Betrouwbare en energie-efficiënte netwerkprotocollen voor draadloze Body Area Networks

Reliable and Energy Efficient Network Protocols for Wireless Body Area Networks

Benoît Latré

Promotoren: prof. dr. ir. I. Moerman, prof. dr. ir. P. Demeester
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Elektrotechniek

Vakgroep Informatietechnologie
Voorzitter: prof. dr. ir. P. Lagasse
Faculteit Ingenieurswetenschappen
Academiejaar 2007 - 2008

UNIVERSITEIT
GENT

ISBN 978-90-8578-217-9
NUR 986
Wettelijk depot: D/2008/10.500/36

Universiteit Gent
Faculteit Ingenieurswetenschappen
Vakgroep Informatietechnologie

Promotoren: Prof. Dr. Ir. Ingrid Moerman
Prof. Dr. Ir. Piet Demeester

Universiteit Gent
Faculteit Ingenieurswetenschappen

Vakgroep Informatietechnologie
Gaston Crommenlaan 8 bus 201, B-9050 Gent, België

Tel.: +32-9-331.49.00
Fax.: +32-9-331.48.99

Dit werk kwam tot stand in het kader van een
specialisatiebeurs van het IWT-Vlaanderen
(Instituut voor de aanmoediging van Innovatie door
Wetenschap en Technologie in Vlaanderen)

Proefschrift tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen:
Elektrotechniek
Academiejaar 2007-2008

Dankwoord

Doctoreren, wat houdt dit nu precies in? Dat is ook de vraag die ik me stelde nu bijna vijf jaar geleden. Ik zat in mijn laatste jaar en deed hier op IBCN mijn thesis. Ergens in het tweede semester polste men "Of ik geen zin had om te doctoreren?" Niet goed wetende wat dit allemaal precies betekende, ging ik er op in. Een keuze die ik me nooit beklaagd heb. Volgens de site van UGent is het behalen van een doctoraat "een bewijs van zelfstandig, origineel wetenschappelijk onderzoek." Dit kan ik alleen maar onderschrijven. Even verder staat echter nog een zeer belangrijke opmerking: "Je staat er zeker niet alleen voor!". Als er iets is dat ik de voorbije jaren geleerd heb, is het dat een doctoraat inderdaad geen solowerk is. Veel mensen hebben immers direct of indirect bijgedragen aan de totstandkoming van dit proefschrift. Zij verdienen het hier vermeld te worden. Aangezien het schier onmogelijk is om iedereen in dit dankwoord op te nemen, wil ik alvast hen bedanken die hier niet vermeld worden.

Als eerste bedank ik onze vakgroepvoorzitter, Prof. Lagasse, voor het ter beschikking stellen van de uitgebreide faciliteiten. Verder wil ik Prof. Demeester, de drijvende kracht achter de explosief groeiende IBCN onderzoeksgroep en copromotor van onderliggend werk, bedanken voor de kansen en de steun die hij me geboden heeft om mijn onderzoek in deze onderzoeksgroep te kunnen uitvoeren. Daarnaast mag ik natuurlijk mijn promotor, Prof. Moerman, niet vergeten voor de ondersteuning, de nodige discussies en vragen waardoor ik mijn werk steeds weer kritisch kon bekijken en nieuwe invalshoeken kon vinden.

Een groot gedeelte van dit werk is tot stand gekomen in samenwerking met Bart Braem van de Universiteit Antwerpen, aan wie ik dan ook grote dank verschuldigd ben. Zonder onze talrijke discussies, ons gezamenlijk programmeer- en publiceerwerk zou het resultaat nooit hetzelfde zijn geweest. Daarnaast mogen ook de mensen niet ontbreken die de propagatie rond het menselijk lichaam bestuderen: Laurens Roelens en later ook Elisabeth Reusens.

Mijn interesse voor draadloze netwerken zou nooit hetzelfde geweest zijn zonder de Mobile-medewerkers van het eerste uur: Liesbeth Peters, Jeroen Hoebeke en Tom Van Leeuwen. Reeds als mijn thesisbegeleiders en later als collega's werkte hun enthousiasme aanstekelijk. Ik was de vierde man in de Mobile groep, die ondertussen uitgegroeid is tot meer dan 10 personen. Een nieuw en uiterst boeiend facet van deze groep is het onderzoek naar sensornetwerken, wat heel nauw aansluit bij onderliggend onderzoek. Ik wil hierbij dan ook de "sensormensen" bedanken voor de goede samenwerking en natuurlijk ook omdat het toffe collega's zijn: Pieter De Mil, Tim Allemeersch, Eli De Poorter, Evy Troublyn en

Lieven Tytgat. Verder nog een speciale vermelding voor Bart Jooris en zijn onvermoeibare inspanningen om het WiLab.t op te starten. Ook de overige Mobile collega's wil ik bij deze bedanken.

Het naarstig werken aan het doctoraat werd een stuk aangenamer door de prettige sfeer onder de bureaugenoten. Waar is de tijd dat we ons eigen appartement hadden in het Zuid, waarvan we op een passende manier afscheid genomen hebben. Bedankt voor de leuke tijd: Didier Colle, Jan Coppens, Philippe De Neve, Andy Van Maele, Jeroen Hoebeke, Matthias Priem, Tom Van Leeuwen en Frederic Van Quickenborne. In de Zuiderpoort kregen we als hechte groep een van de felbegeerde hoekbureaus toegewezen. Zo konden we dankzij ons zicht op het viaduct en op Ledeberg de nodige inspiratie vinden indien nodig. In de loop der jaren werd ons bureau aangevuld met Dries Naudts, Stefan Bouckaert, Nicolas Staelens en sinds kort ook met Irina Ballan, David Carels en Willem Vereecken. Ook zij zijn bij deze bedankt!

Gedurende de voorbije jaren heb ik altijd kunnen rekenen op de technische support van ons onovertroffen admin-team: Wouter Adem, Bert De Vuyst, Pascal Vandeputte en Brecht Vermeulen. Voor de ondersteuning bij de practica stond Andy Van Maele altijd paraat. Dankzij de mensen van de administratie raakte ik niet bedolven onder al het papierwerk, werden mijn vliegtickets en hotels keurig geboekt en werden mijn reisonkosten correct terugbetaald: bedankt Martine, Ilse, Davinia, Karien, Marleen en Bernadette.

Gedurende de zware inspanningen, frustraties van lastige implementaties en naderende deadlines, werd er ook gezorgd voor de nodige ontspanning. Hierbij denk ik aan de BBQ's, de spelletjesavonden, de vrijdagavond apero's, de fuiven die we onveilig gemaakt hebben, de IBCN- (mini) weekendjes waar te luid Vlaamse schlagers zingen blijkbaar niet door iedereen geapprecieerd werd, de quizzen de maandagavond,... Bedankt voor de vele leuke momenten en de vriendschap van de voorbije jaren: Bart D., Bart L., Bert, Bruno, Erik, Jan VO, Jurgen, Koert, Masha, Niels, Pieter, Sarah, Sofie DM, Sofie VH, Stefan en Stijn.

Daarnaast waren er ook de mensen buiten het werk die voor de nodige ontspanning hebben gezorgd tijdens de weekendjes in Vlessart, de zoektocht naar oliebollen in Zwijnaarde, de oudejaarsavonden, de verjaardagsdrinks,... De "bende" stond altijd paraat! Bedankt Annelies & Joost, Evelyne & Jody, Karen & Bert en Katrien, Kurt, Liam & ?.

Verder zou ik ook graag mijn ouders en broer bedanken voor de hun onvoorwaardelijke steun van de laatste jaren. Ook een bedankje voor mijn toekomstige schoonzus Joke die de moed heeft gehad mijn thesis na te lezen op spellingsfoutjes.

Nu rest me slechts nog 1 speciaal iemand te bedanken die me zeer na aan het hart ligt: Lisa. Bedankt voor de lieve steun en het luisterend oor het voorbije anderhalf jaar en zeker op de momenten dat het me allemaal wat teveel werd. Je gaf me de steun die ik nodig had. Zonder jou zou ik nooit staan waar ik nu sta. Samen komen we er wel!

Gent, juni 2008
Benoît Latré

Table of Contents

List of Figures

List of Tables

List of Symbols and Acronyms

Symbols

α_i	Receiving period for node i, expressed in slots
α_i^A	α_i when aggregation is used
β_i	Waiting period for child node i, expressed in slots
β_i^A	β_i when aggregation is used
$\gamma(i)$	Total waiting period for a node i
δ_i	Slots needed by node i to send own data in one cycle
Δ	Duty cycle
ϵ	Ratio between slot length in data subcycle and control subcycle
ζ	Number of children in each branch of the tree
η	Path loss exponent
κ_i	Length of a packet from node i in bits
λ	wave-length
ϕ	Aggregation factor
ρ_i	Sleep ratio of node i
σ	Standard deviation of the path loss
A_i	The set of edges on the path from node i to the personal device
Ch_n	Set of children of node n
$CP_{i,j}$	Connection probability of the path (i,j)
d_0	Reference distance
d_{char}	Characteristic distance
E_{amp}	Dynamic transmitter energy
E_{RXelec}	Static receiver energy
E_{TXelec}	Static transmitter energy
E_{MH}	Energy consumption in multi-hop topology
E_{SH}	Energy consumption in single-hop topology
E_R	Energy consumption by a relay device
k	Number of bits transmitted
L	Number of levels in the tree

$\log_{10}$	Logarithm in base 10
N	Number of nodes in the network
p	Probability for link connectivity
$P^j_{r,dB}$	Received signal strength from node j
$P^j_{s,dB}$	Transmitted signal strength from node j
PL_{dB}	Path loss, expressed in dB
$PL_{0,dB}$	Path loss at reference distance d_0, expressed in dB
S	Sink or personal device
$t_{arrival}$	Packet inter arrival rate
T_{CC}	Length of the control cycle
T_{cycle}	Length of one cycle
T_{on}	Time the radio is on
T_{off}	Time the radio is switched off
T_{WC}	Length of a WASP-cycle
TS_i	Number of time slots needed by node i
V	Set of all nodes
V_n	Set of nodes in the tree below node n
V_r	Set of relay devices in the network

A

ACK	Acknowledgment
API	Application Programming Interface

B

BAN	Body Area Network
BASN	Body Area Sensor Network
BCC	Body-Coupled Communication
BCU	Body Control Unit
BE	Back off Exponent
BER	Bit Error Rate
BFS	breadth-first search
BPSK	Binary Phase Shift Keying
BO	Back Off
BSN	Body Sensor Network
BWE	BandWidth Efficiency

C

CAP	Contention Access Period
CBR	Constant Bit Rate
CDMA	Code Division Multiple Access
CFP	Contention Free Period
CICADA	Cascading Information Retrieval by Controlling Access with Distributed Slot Assignment
CICADA-S	CICADA with Security
CP	Connection Probability
CPU	Control Processing Unit
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CVD	Cardiovascular Disease
CTR	Counter

D

DC	Duty Cycle

E

ECG	Electrocardiogram
EEG	Electroencephalogram
EM	Elektromagnetic
EMG	Electromyogram

F

FCC	Federal Communications Commission
FTR	Footer

G

GP	General Practitioner

GPRS	General Packet Radio Service
GTS	Guaranteed Time Slot

H

HDR	Header
HF	High Frequency

I

ID	IDentity
IEEE	Institute of Electrical and Electronics Engineers, Inc.
IFS	Inter Frame Space
ILP	Integer Linear Programming
IM3	Interactive Mobile Medical Monitoring
IP	Internet Protocol
ISM	Industrial Scientific Medical band

L

L2	Layer 2 (MAC layer)
L3	Layer 3 (Network layer)
LAN	Local Area Network
LEACH	Low Energy Adaptive Clustering Hierarchy
LF	Low Frequency
LIFS	Long IFS
LOS	Line Of Sight
LTR	Least Temperature Routing

M

MAC	Medium Access Control
MH	Multi-hop
MOFBAN	Modular Framework for BAN
MPDU	Mac Protocol Data Unit

N

NLOS	Non Line Of Sight
NS-2	Network Similator 2

O

OSI	ISO reference model for Open Systems Interconnection

P

PAN	Personal Area Network
PD	Personal Device
PDA	Personal Digital Assistant
PDU	Protocol Data Unit
PEGASIS	Power Efficient Gathering in Sensor Information System
PER	Packet Error rate
PHY	PHYsical layer
PL	Path Loss
PLL	Phase Locked Loop

Q

QoE	Quality of Experience
QoL	Quality of Life
QoS	Quality of Service

R

RF	Radio Frequency
RTP	Real-time Transport Protocol

RTT	Round Trip Time
Rx	Receiver

S

S-MAC	Sensor MAC
SAR	Specific Absorption Rate
SH	Single-hop
SIFS	Short IFS
SNA	Sensor Network Architecture
SNR	Signal to Noise Ratio
SP	Silence Period

T

T	Time
TARA	Thermal Aware Routing Algorithm
TCP	Transport Control Protocol
TDMA	Time Division Multiple Access
TEG	Thermoelectric Generator
TP	Throughput
TPE	Throughput Efficiency
TR	Technical Report
Tx	Transmitter

U

UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
UWB	Ultra Wide Band

V

VCO	Voltage Controlled Oscillator

W

WAN	Wide Area Network
WASP	Wireless Autonomous Spanning Tree
WiMax	Worldwide interoperability for Microwave ACCess
WiFi	Wireless Fidelity
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Network
WSN	Wireless Sensor Network
WSAN	Wireless Sensor and Actuator Network

Nederlandse samenvatting –Summary in Dutch–

De voorbije jaren zijn elektronische apparaten alsmaar kleiner geworden terwijl ze steeds complexere taken aankunnen. Het is nu zelfs mogelijk om sensoren te ontwikkelen die zo klein zijn dat ze gemakkelijk en zonder veel hinder in de kledij, op het lichaam of onder de huid van een persoon geplaatst kunnen worden. Deze sensoren zijn bovendien uitgerust met een radio waardoor ze draadloos met elkaar kunnen communiceren. Op deze manier wordt een draadloos netwerk gecreëerd dat een Wireless Body Area Network (WBAN) wordt genoemd. Het draadloze karakter en de grote verscheidenheid aan sensoren bieden uitzicht op een brede waaier van praktische en innovatieve toepassingen. Het meest voor de hand liggend zijn de medische toepassingen, waarbij bepaalde lichaamsfuncties (EKG, EEG, bloeddruk, temperatuur, ...) van een patiënt kunnen gemonitord worden terwijl de patiënt over zijn volledige bewegingsvrijheid blijft beschikken. Dit bevordert niet alleen de levenskwaliteit van de patiënt, maar laat ook een nauwkeurigere diagnose toe. Zo kan een sensor bijvoorbeeld het glucosegehalte in het bloed meten. Deze informatie wordt dan draadloos verzonden naar een ander apparaatje dat insuline toedient indien nodig. Daarnaast is ook de sportwereld geïnteresseerd om atleten zowel op technisch als op conditioneel vlak beter te kunnen begeleiden. Een WBAN heeft ook veel te bieden voor multimediatoepassingen. Denken we maar aan draadloze MP3-spelers, camera's, minibeeldschermen en bewegingssensoren voor een virtual reality omgeving.

In een WBAN worden verschillende soorten apparaten (ook wel nodes genoemd) gebruikt die allemaal uitgerust zijn met een radio. De meest voorkomende zijn de sensoren en actuatoren. Dit zijn apparaten die respectievelijk fysische parameters meten ofwel één of andere actie ondernemen, zoals het toedienen van medicijnen. Deze apparaten staan via een WBAN draadloos in verbinding met een intelligent apparaat, zoals een PDA of een mobiele telefoon. Dit apparaat wordt een centrale node genoemd en staat in voor de communicatie met de buitenwereld. Zo kunnen de verkregen metingen worden doorgestuurd naar een arts of het ziekenhuis. Als laatste type zijn er nog de relay-apparaten. Deze beschikken niet over een sensor of een actuator, maar bestaan enkel uit een radio met netwerkfunctionaliteit. Zij ontvangen de data afkomstig van de ene node en sturen die door naar de andere node. Op deze manier wordt communicatie mogelijk tussen twee nodes die elkaar niet kunnen bereiken. De nodes in een WBAN zijn bovendien heterogeen van nature. Ze hebben verschillende eigenschappen zoals onder andere een hoge of lage doorvoersnelheid, een kleine of een grote batterij of al dan niet kritische data.

Een WBAN stelt dan ook enkele specifieke eisen aan het netwerk. Er wordt gewerkt met zeer kleine apparaatjes waarbij er maar een beperkte ruimte voorzien is voor een batterij. Het is dus zeer belangrijk om het energieverbruik zoveel mogelijk te beperken. Dit kan door een zo zuinig mogelijke radio te bouwen, maar ook door het ontwikkelen van energie-efficiënte netwerkprotocollen die bepalen hoe de communicatie tussen de radio's verloopt. Een mogelijke techniek is het verminderen van het zendvermogen. Dit zorgt er echter voor dat de sensoren zullen moeten samenwerken om de data naar de centrale node te sturen. De data wordt van de ene node naar de andere gezonden tot de centrale node bereikt is. Dit wordt multi-hop communicatie genoemd. Een bijkomend voordeel is dat op deze manier ook de invloed van de elektromagnetische stralingen op het lichaam beperkt wordt. Niet alleen het energieverbruik is echter relevant, ook de betrouwbaarheid van de communicatie is zeer belangrijk. De data die verstuurd wordt bevat immers medische informatie waarbij men er zeker van moet zijn dat de data ook effectief ontvangen wordt. Het mag bijvoorbeeld niet gebeuren dat een kritisch bericht verloren gaat. Bovendien moet een WBAN de heterogeniteit van de nodes ondersteunen.

In dit proefschrift wordt de nadruk gelegd op de ontwikkeling van energie-efficiënte en betrouwbare netwerkprotocollen voor WBANs. Hierbij zal ondermeer gebruik gemaakt worden van multi-hop routering en een verbeterde interactie tussen de verschillende netwerklagen. Verschillende mechanismes die het energieverbruik laten dalen en die de betrouwbaarheid verhogen worden voorgesteld.

In eerste instantie worden de fysische kenmerken van de communicatie bestudeerd. Er is een model opgesteld voor de evolutie van het padverlies bij propagatie van elektromagnetische golven langs het lichaam. Hierbij is het belangrijk te noteren dat dit padverlies veel groter is dan bij directe communicatie in de vrije ruimte. Er moet bovendien rekening gehouden worden met de kromming van het lichaam. De communicatie vanaf de voorkant van het lichaam naar de achterkant ondervindt zeer hoge verliezen waardoor een direct pad tussen de voor- en achterkant quasi onmogelijk is. Dit model is vervolgens gebruikt om de kans op connectiviteit tussen de nodes te bepalen. In plaats van een cirkelvormige voorstelling van het bereik van de radio (alles binnen een straal r is bereikbaar), wordt gebruik gemaakt van een lognormale voorstelling. Deze geeft een realistischer beeld van de topologie van het netwerk. Vervolgens wordt een energiemodel voorgesteld voor een radio in een WBAN op basis van drie bestaande commerciële radio's.

Het volgende hoofdstuk gebruikt de ontwikkelde modellen om te onderzoeken welke netwerktopologieën energie-efficiënt en betrouwbaar beschouwd kunnen worden. Aan de hand van een lijntopologie en een boomstructuur wordt de invloed nagegaan van het verlagen van het zendvermogen, het gebruik van multi-hop communicatie (door de nodes te laten samenwerken) en het aggregeren of samenvoegen van data naar de centrale node. Vervolgens worden ook relay-apparaten toegevoegd aan het netwerk. Met behulp van ILP (integer linear programming) wordt de optimale oplossing nagegaan. Door de nodes te laten samenwerken bij het doorsturen van de data en door extra relay nodes toe te voegen aan het netwerk, wordt een energie-efficiënter netwerk bekomen. Voor grotere netwerken is er zelfs een verbetering merkbaar van 30%. Ook de betrouwbaarheid van het netwerk wordt verhoogd wanneer multi-hop communicatie wordt gebruikt, terwijl het energieverbruik maar lichtjes toeneemt.

De verschillende toepassingen van een WBAN vragen om specifiek aangepaste maatregelen of protocollen, zoals bijvoorbeeld het al dan niet gebruiken van data-aggregatie, aanpassen van het zendvermogen en meer of minder betrouwbaarheid. Voor elke netwerklaag moet er dan bepaald worden welk protocol het meest geschikt is en voor elk van die protocollen moeten de meest optimale instellingen genomen worden. Om dit proces te vereenvoudigen wordt een structuur voorgesteld voor een netwerkarchitectuur, MOFBAN genaamd. MOFBAN maakt het mogelijk om op een eenvoudige manier netwerkprotocollen toe te voegen aan het netwerk of ze aan te passen. MOFBAN is modulair opgebouwd, d.w.z. dat alle functionaliteit ondergebracht is in modules. Op deze manier wordt dubbel gebruik van enkele functies (zoals een foutcontrole zowel op de MAC-laag als netwerklaag én transportlaag) vermeden. Verder kan ervoor gekozen worden om de verschillende modules al dan niet te gebruiken en dit naargelang de behoefte van de applicatie of de mogelijkheden van de node. Op deze manier wordt de heterogeniteit van de nodes en applicaties ondersteund. Om communicatie tussen de verschillende modules en nodes mogelijk te maken wordt een modulaire pakketstructuur voorgesteld.

Na het vastleggen van de netwerkarchitectuur wordt in een volgend hoofdstuk overgegaan op de studie van de netwerkprotocollen voor WBANs. In eerste instantie worden de theoretische maximale doorvoersnelheid (throughput) en de vertraging van IEEE 802.15.4 analytisch bepaald. IEEE 802.15.4 is een MAC-protocol speciaal ontwikkeld voor sensornetwerken. Toch blijkt dat deze niet geschikt is voor het gebruik in een WBAN wegens een te hoog energieverbruik. Vervolgens wordt WASP voorgesteld, een cross-layer protocol dat zowel de MAC-laag als de routeringslaag in een WBAN voor zijn rekening neemt. Het protocol zet een boomstructuur op en verdeelt de tijdsas in stukken van dezelfde lengte, slots genaamd. Het toekennen van deze slots gebeurt op een gedistribueerde wijze: elke node zendt een WASP-schema naar elk van zijn kinderen waarbij hen verteld wordt wanneer ze mogen zenden.

In een volgend hoofdstuk wordt CICADA voorgesteld, een netwerkprotocol dat de tekortkomingen van WASP oplost. Het protocol zet net als WASP een boomstructuur op. De tijdsas wordt nu echter ook verdeeld in een controlegedeelte en een datagedeelte die na elkaar komen en elk op hun beurt ook uit slots bestaan. Het controlegedeelte wordt gebruikt voor het doorsturen van de schema's die de nodes vertellen wanneer ze mogen zenden in het datagedeelte. Het verkeer gaat van de centrale nodes naar de andere nodes. In het datagedeelte worden de datapakketten verstuurd van de nodes naar de centrale node. Op deze manier wordt het mogelijk om bidirectioneel verkeer te ondersteunen en om de vertraging zo laag mogelijk te houden. CICADA wordt analytisch geëvalueerd wat betreft het energieverbruik en de throughput waaruit blijkt dat CICADA weinig energie verbruikt maar toch over een grote throughput beschikt. Het protocol is geïmplementeerd in NS-2 en vergeleken met een bestaand MAC-protocol voor sensornetwerken. CICADA blijkt energie-efficiënter en betrouwbaarder te zijn. Als slot worden er extra mechanismen toegevoegd voor de betrouwbaarheid en de beveiliging van het dataverkeer.

Het laatste hoofdstuk tenslotte bevat een overzicht van de voornaamste conclusies.

English Summary

In the last few years, advances in electronics have enabled the development of small yet intelligent devices. These tiny devices can easily be attached to clothing, placed on the body or even implanted inside the human body where they can measure physiological parameters. When the sensors are equipped with a wireless interface, a new type of network spanning the entire body takes form: a Wireless Body Area Network (WBAN). The wireless nature of the network and the wide variety of sensors offers numerous new, practical and innovative applications. A motivating example can be found in the world of health monitoring. The sensors of the WBAN for example measure the heartbeat, the body temperature or record a prolonged electrocardiogram (ECG). Using a WBAN, the patient experiences a greater physical mobility and is no longer compelled to stay in a hospital. This not only improves the patient's Quality of Live, but also allows a more accurate diagnosis. As an example, a sensor can measure the glucose level in the blood. The obtained data is then wirelessly transmitted to another device that administers insulin if needed. A WBAN has also gained the attention of the world of sports so as to provide a better guidance of athletes. Further, a WBAN can offer solutions for multimedia applications such as wireless MP3-players, small screens and motion sensors to create an environment for virtual reality.

A WBAN is made up out of several kinds of devices (also called nodes), all equipped with a wireless interface. The most common devices are the sensors and the actuators. The former measure physiological data while the latter can undertake a specific action according to the data they receive from the sensors or through interaction with the user, such as administering medications. The WBAN provides a wireless connection between these devices and a more intelligent device, e.g. a PDA or a smartphone. This device acts as a sink for the sensor information and is called a personal device. It is responsible for the connection with other networks. For example, the obtained data can be forwarded to a General Practitioner or a hospital. A last type of devices are the relay devices. They do not have any sensing or actuating capabilities, but are only equipped with a radio and network functionality. The received data is forwarded from one node to another. Doing so, a relay device enables the communication between two nodes that are placed too far apart. Furthermore, the nodes in a WBAN are heterogeneous as they have various requirements such as low or high throughput, small or large battery capacity or (non-) critical data.

It is clear that a WBAN imposes the networks some strict and specific requirements. The devices are tiny, leaving only limited space for a battery. It is therefore

of uttermost importance to restrict the energy consumption in the network. A possible solution is to build a low power radio, but also the development of energy efficient protocols that regulate the communication between the radios. A conceivable technique is lowering the transmit power of the radios. However, this causes that the sensors will have to cooperate in order to send the data to the personal device. The data is sent from one device to another, a process called multi-hop communication. An extra advantage is that in this way also the impact of the electromagnetic waves on the body is restricted. Besides the energy consumption, it is also important to consider the reliability of the communication. The data sent contains medical information and one has to make sure that the data is correctly received at the personal device. It is not allowed that a critical message is lost. In addition, a WBAN has to support the heterogeneity of its devices.

This thesis focuses on the development of energy efficient and reliable network protocols for WBANs. Considered solutions are the use of multi-hop communication and the improved interaction between the different network layers. Mechanisms to reduce the energy consumption and to grade up the reliability of the communication are presented.

In a first step, the physical layer of the communication near the human body is studied and investigated. A path loss model for the propagation of electromagnetic waves along the body is presented. It is important to notice that this path loss is considerably higher than the path loss obtained in free space. Additionally, one has to take into account the curvature of the body. Communication between the front and the back of the body experiences extremely high path losses, making direct communication between the front and the back virtually impossible. This model is subsequently used to model the probability of a connection between two nodes on the body. Instead of assuming a circular coverage area (all nodes within a distance r can be reached), a lognormal distribution is used to account for the shadowing effects. This gives a more realistic view of the network topology. Further, an energy model for the radio in WBANs is proposed. This energy model considers the receiver and transmitter energy of the radio. The transmitter energy contains two parts: a static part and a part for the transmitter amplifier that scales with the distance. Based on the previously proposed path loss model, the parameters for three frequently used radios are determined.

In the following chapter, the connectivity model and energy model for the radio are used to investigate which network topologies can be considered as the most energy efficient and reliable. By means of a line topology and a tree network, the effect of lowering the transmit power, using multi-hop communication (by letting the nodes cooperate) and aggregating data to the personal device are investigated. In the next step relay devices are added to the network. An ILP (integer linear programming) formulation is drawn up to work out the optimal solution. By letting the nodes cooperate and by adding additional relay devices near the personal device, a more energy efficient network is obtained. For larger networks a lifetime increase of more than 30% was found. The influence of reliability on the network topology has been evaluated by setting up an ILP. The use of multi-hop communication improves the reliability. Overall, it was concluded that a multi-hop architecture is the best choice for a WBAN and that one has to deal with a trade-off between energy efficiency and reliability. Adding relay devices is helpful for both the energy efficiency and reliability.

The different applications in a WBAN (i.e. medical monitoring, sport monitoring, gaming etc.) require specific measures like data-aggregation, changing the transmitting power and more or less reliability. For each network layer the most suitable protocol needs to be selected and each protocol must be tuned to the most optimal settings. In order to ease this process we present MOFBAN, a lightweight framework for a network architecture in WBANs. MOFBAN enables an easier way to add or adjust network protocols. It has a modular structure, i.e. all functionality is implemented in modules. Doing so, duplication of functionality can be avoided, for example error correction that happens on the MAC-layer, network layer and transport layer. Depending on the capabilities of the node and/or the application, more modules (and thus network functionality) can be added. This way, heterogeneous networks can be supported. So as to enable communication between the nodes and the modules, a modular header structure is presented.

Once the network architecture is given, we study the network protocols for WBANs in the next chapter. In particular, the maximum throughput and minimum delay of IEEE 802.15.4 are determined. Although IEEE 802.15.4 was specifically developed as MAC-protocol for wireless sensor networks, it is not suitable for use in WBANs because of a too high energy consumption. Then a new cross layer protocol for wireless Body Area Networks that both handles channel medium access and routing is presented: WASP. The protocol sets up a spanning tree and divides the time axis in slots. The slot allocation is done in a distributed manner. Every node sends out a proprietary WASP-scheme to its children to inform them of the following level when they are allowed to send.

The following chapter introduces CICADA, a network protocol that solves WASP's shortcomings. Like WASP, the protocol sets up a spanning tree. The time axis is now divided into a control part and a data part, each containing several time slots. The control part is used to send the schemes from the parent nodes to the children nodes. These schemes tell the children when they are allowed to send during the data part. In the data part, the data packets are sent from the nodes to the personal device. Doing so, bidirectional traffic is supported and the delay is kept to a minimum. CICADA's energy consumption and the throughput are evaluated analytically. CICADA proves to consume little energy and has a high throughput. The protocol is implemented in NS-2 and compared with S-MAC, an existing and much used protocol for sensor networks. CICADA turned out to be more energy efficient and to experience less packet loss. Finally additional mechanisms are added to further support the reliability and to provide security.

The last chapter contains an overview of the most important contributions and conclusions.

1
Introduction

This chapter provides a definition of Wireless Body Area Networks and a small overview of its application areas. The main problems and challenges are pointed out. Further we describe the main contributions of this research and an outline of this work is given. We conclude with an overview of manuscripts published.

1.1 Wireless Body Area Networks

The aging population in many developed countries and the rising costs of health care has triggered the introduction of novel technology-driven enhancements to current health care practices. Recent advances in electronics have enabled the development of small and intelligent (bio) medical sensors which can be worn on or implanted in the human body. These sensors need to send their data to an external medical server where it can be analyzed and stored. Using a wired connection for this purpose turns out to be too cumbersome and involves a high cost for deployment and maintenance. It is more cost efficient to equip the sensors in and on the body with a wireless interface to enable an easier application [1]. The patient experiences a greater physical mobility and is no longer compelled to stay in a hospital. This process can be considered as the next step in enhancing the personal health care and in coping with the costs of the health care system. As where eHealth is defined as the health care practice supported by electronic processes and communication, the health care is now going a step further by becoming mobile. This is referred to as mHealth [2, 3]. In order to fully exploit the benefits of wireless tech-

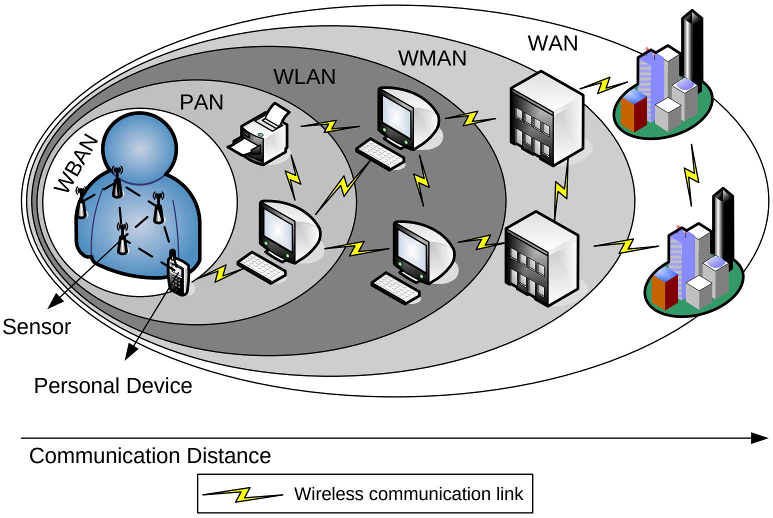


Figure 1.1: Positioning of a Wireless Body Area Network in the realm of wireless networks.

nologies in telemedicine and mHealth, a new type of wireless area network needs to be developed: a wireless on-body network or a Wireless Body Area Network (WBAN). This term was first coined by Van Dam et al. in 2001 [4] and received the interest of several researchers [5–9].

A Wireless Body Area Network provides continuous health monitoring and real-time feedback to the user and the medical personnel. Furthermore, the measurements can be recorded over a longer period of time, improving the quality of the measured data [10]. Interaction with the user or other persons is usually handled by a personal device, e.g. a PDA or a smartphone which acts as a sink for wireless sensor information. Generally speaking, one can distinguish two types of devices: sensors and actuators. The sensors are used to measure certain parameters of the human body, either externally or internally. Examples include measuring the heartbeat, body temperature or recording a prolonged electrocardiogram (ECG). The actuators or actors on the other hand take some specific actions according to the data they receive from the sensors or through interaction with the user, e.g. an actuator equipped with a built-in reservoir and pump for administering the correct dose of insulin to diabetics based on the measurements of glucose level.

In Figure 1.1, a WBAN is compared with other types of wireless networks, such as Personal (PAN), Wireless Local (WLAN), Wireless Metropolitan (WMAN) and Wide Area Networks (WAN). These types of networks can be considered as an enabling technology of an ad hoc network in general [6]. A WBAN is operated close to the human body and its communication range will be restricted within a few meters, with typical values around 1-2 meters. While a WBAN is devoted to interconnection of one-person's wearable devices, a PAN is a network in the environment around the person. The communication range can reach up to 10 meters for high data rate applications and up to several 10 meters for low data rate applications. A WLAN has a typical communication range up to several 100 meters. Each type of network has its typical enabling technology, defined by the IEEE. A PAN uses IEEE 802.15.1 (Bluetooth) or IEEE 802.15.4 (ZigBee), a WLAN uses IEEE 802.11 (WiFi) and a WMAN IEEE 802.16 (WiMax). The communication in a WAN can be established via satellite links. Recently, the IEEE has set up a working group to work out a standard for communication in a WBAN (IEEE 802.15.6) [11]

When referring to a WBAN where each node comprises a biosensor or a medical device with sensing unit, some researchers use the name Body Area Sensor Network (BASN) or in short Body Sensor Network (BSN) instead of WBAN. These networks are very similar to each other and share the same challenges and properties. In this thesis, we will use the term WBAN which is also used by the IEEE.

1.1.1 A Motivating Application: Patient Monitoring

In the last few years, wearable monitoring systems have gained the attention of various researchers, especially in the area of health care. An important task of such a system is to collect physiological parameters like the heartbeat, body temperature etc. To this purpose, several sensors are placed in clothes, directly on the body or under the skin of a person. These measure the temperature, blood pressure, heart rate, ECG, EEG, respiration rate, SpO_2-levels etc. An example of a medical WBAN is shown in Figure 1.2. As the WBAN allows continuous monitoring of these parameters, whether the patient is in the hospital, at home or on the move, the patient will no longer need to keep to his bed, but will be able to move around freely. Furthermore, the data obtained in a large time interval in the patient's natural environment offers a clearer view to the doctors than data obtained during short stays at the hospital [10].

Next to sensing devices, the patient has actuators which act as drug delivery systems. The medicine can be delivered on predetermined moments, triggered by an external source (i.e. a doctor who analyzes the data) or immediately when a sensor notices a problem. One example is the monitoring of the glucose level in the blood of diabetics. If the sensor monitors a sudden drop of glucose, a signal can be sent to the actuator in order to start the injection of insulin. Consequently, the patient will experience fewer nuisances from his disease. Another example of an actuator is a spinal cord stimulator implanted in the body for long-term pain relief [12].

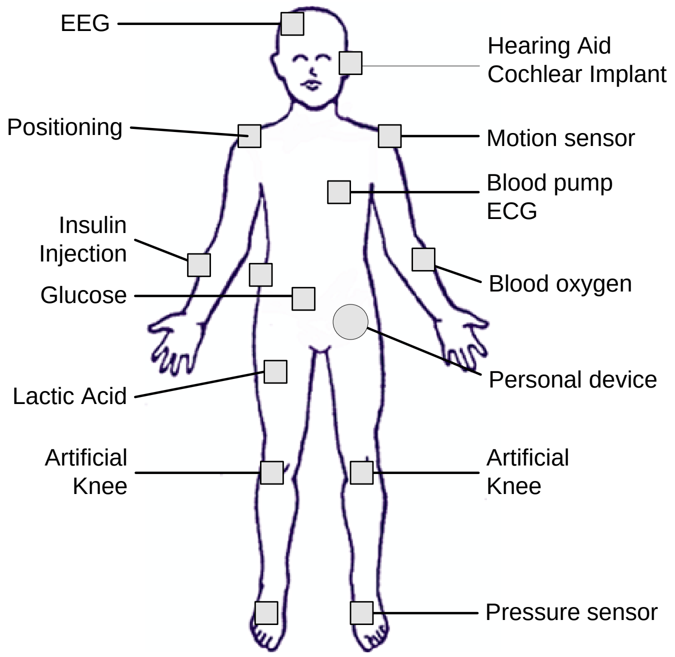


Figure 1.2: Example of patient monitoring in a Body Area Network.

Worldwide, more than 246 million people suffer from diabetes, a number that is expected to rise to 380 million by 2025 [13]. Frequent monitoring enables proper dosing and reduces the risk of fainting and in later life blindness, loss of circulation and other complications [13]. A WBAN can prove to be helpful in this case. Another application for a WBAN is cardiac monitoring. Cardiovascular disease (CVD) is the main cause of death in the world (30% of all death). Worldwide, about 17.5 million people die of heart attacks or strokes each year; in 2015, almost 20 million people will die from CVD [14]. These deaths can often be prevented with proper health care.

A WBAN can also be used to offer assistance to the disabled. For example, a paraplegic can be equipped with sensors determining the position of the legs or with sensors attached to the nerves [15]. In addition, actuators positioned on the legs can stimulate the muscles. Interaction between the data from the sensors and the actuators makes it possible to restore the ability to move. Another example is aid for the visually impaired [16, 17]. An artificial retina, consisting of a matrix of micro sensors, can be implanted into the eye beneath the surface of the retina. The artificial retina translates the electrical impulses into neurological signals. The input can be obtained locally from light sensitive sensors or by an external camera mounted on a pair of glasses.

Other applications of a WBAN are, amongst others [10, 18–20]:

- Elderly people will be able to live longer in their own home and will not need to move to an expensive rest home or at least not until a later age; sensors can monitor the health conditions or the behavior, e.g. a fall;
- The prevention of sudden infant death syndrome ;
- Stress monitoring by recording the blood pressure, heart rate etc.;
- Sensors placed in the nasal area or the tongue may be used to detect harmful toxins in for example the food ingested and air inhaled etc;
- The introduction of a BAN enables to tune more effectively the training schedules of professional athletes.

The devices used in a WBAN are becoming tinier, even less than 1 cm^3 [21, 22]. The ultimate goal is to create a small and smart band-aid containing all the necessary technology for sensing and communication with a base station, for example the Sensium-platform [23]. This will allow an even easier application and acceptance of WBANs.

These examples illustrate the need for continuous monitoring and the usefulness of WBANs. However there are other examples of diseases that would benefit from continuous or prolonged monitoring, such as hypertension, asthma, Alzheimer's disease, Parkinson's disease, renal failure, post-operative monitoring etc. These applications are an indicator for the size of the market for WBANs. The number of people suffering from diabetics or CVD and the growing percentage of people in the population age 60 years and older will grow in the future. Even without any further increase in world population by 2025 this would mean a very large number of potential customers. Devices offering support to an aging population could use WBAN technology to provide the connectivity functionality to support the elderly in managing their daily life and medical conditions [20].

The developers of WBANs will have to take into account the privacy issues. After all, a WBAN can be considered as a potential threat to freedom, if the applications go beyond "secure" medical usage, leading to a Big Brother society. Social acceptance would be the key to this technology finding a wider application. Therefore, considerable effort should be put in securing the communication and making sure that only authorized persons can access the data.

Another area of application can be found in the domain of public safety where the WBAN can be used by firefighters or policemen. The WBAN monitors for example the level of toxics in the air and warns the firefighters if a life threatening level of a toxic is detected. As a last application we mention the use of the network in a military environment [24]. The WBAN can be used for stress monitoring and for detection of a biological or chemical attack.

Next to purely medical applications, a WBAN can include appliances such as an MP3-player, head-mounted (computer) displays, a microphone, a camera, advanced human computer interface such as a neural interface etc [15]. As such, the WBAN can also be used for gaming purposes and in virtual reality.

This overview only offers a small cross section of possible applications. Many more applications exist for WBANs. The main characteristic of all these applications is that WBANs improve the user's Quality of Life. In the next section, we give an overview of existing projects that use a WBAN.

1.1.2 Existing Projects

To date, the research in the area of WBANs mainly focuses on building a system architecture and service platform and in lesser extent on developing networking protocols. In the following, we will give a non-exhaustive list of existing projects for WBANs.

Otto *et al.* [8] and Jovanov *et al.* [25] present a system architecture of a wireless body area sensor network which both handles the communication within the WBAN and between the WBANs and a medical server in a multi-tier telemedicine system. The communication between the sensors and the sink is single-hop, slotted and uses ZigBee or Bluetooth. The slots are synchronized using beacons periodically sent by the sink. They use off-the-shelf wireless sensors to design a prototype WBAN such as the Tmote sky platform from Moteiv [26].

The Tmote sky platform is also used in the CodeBlue-project [27, 28] where WBANs are used in rapid disaster response scenarios. A wearable computer attached to the patients wrist, i.e. a Tmote Sky mote, forms an ad hoc wireless network with a portable tablet PC. They developed a wireless two-lead ECG, a wireless pulse oximeter sensor and a wireless electromyogram (EMG).

Ayushman is a sensor network based medical monitoring infrastructure that can collect, query and analyze patient health information in real-time [29]. A wireless ECG, gait monitoring and environment monitoring was developed using off-the-shelf components with a Mica2 wireless transceiver. They have further developed the necessary software for consulting the data at a remote client.

The Human++ project by IMEC-NL [22] aims *"to achieve highly miniaturized and autonomous sensor systems that enable people to carry their personal body area network."*. They have developed an ambulatory EEG/ECG system with a transmitter working on 2.4 GHz. This system can run for approximately 3 months using 2 AA batteries. In order to obtain a longer autonomy, the project also investigates energy scavenging with thermoelectric generators (TEG). In 2006, a wireless pulse oximeter was presented, fully powered by the patient's body heat. Further, the project investigates new wireless technologies such as UWB so as to make an ultra-low power transmitter.

The European funded Mobihealth project [30] provides a complete end-to-end mhealth platform for ambulant patient monitoring, deployed over UMTS and GPRS networks. The MobiHealth patient/user is equipped with different sensors that constantly monitor vital signals, e.g. blood pressure, heart rate and electrocardiogram (ECG). Communication between the sensors and the personal device uses Bluetooth or ZigBee and is single hop.

The German funded BASUMA-project (Body Area System for Ubiquitous Multimedia Applications) [31] aims at developing a full platform for WBANs. As communication technique, a UWB-frontend is used and a MAC-protocol based on IEEE 802.15.3. This protocol also uses time frames divided into contention free periods (with time slots) and contention access periods (CSMA/CA).

The Flemish IBBT IM3-project (Interactive Mobile Medical Monitoring) focuses on the research and implementation of a wearable system for health monitoring [32]. Patient data is collected using a WBAN and analyzed at the medical hub worn by the patient. If an event (e.g. heart rhythm problems) is detected, a signal is sent to a health care practitioner who can view and analyze the patient data remotely.

In November 2007, the IEEE 802.15 Working Group has started a task group for WBANs: IEEE 802.15.6 [11]. This task group will *"develop a communication standard optimized for low power devices and operation on, in or around the human body (but not limited to humans) to serve a variety of applications including medical, consumer electronics/personal entertainment and other"*. The creation of this group stresses the importance of the research with respect to WBANs.

1.2 Problems and Challenges

A WBAN has some typical properties in comparison with other types of networks such as ad hoc networks and sensor networks:

- The devices used have limited energy resources available as they have a very small form factor (often less than 1 cm^3). Furthermore, for most devices it is impossible to recharge or change the batteries although a long lifetime of the device is wanted (up to several years or even decades for implanted devices). Hence, the energy resources and consequently the computational power and available memory of such devices will be limited;
- The network consists of devices in and on the body which are in each others vicinity and hence may interfere strongly;
- A WBAN only has a limited number of devices (less than 50) where each device is equally important and devices are only added when they are needed for an application (i.e. no redundant devices);
- An extremely low transmit power per node (noninvasive) to minimize interference and cope with health concerns;

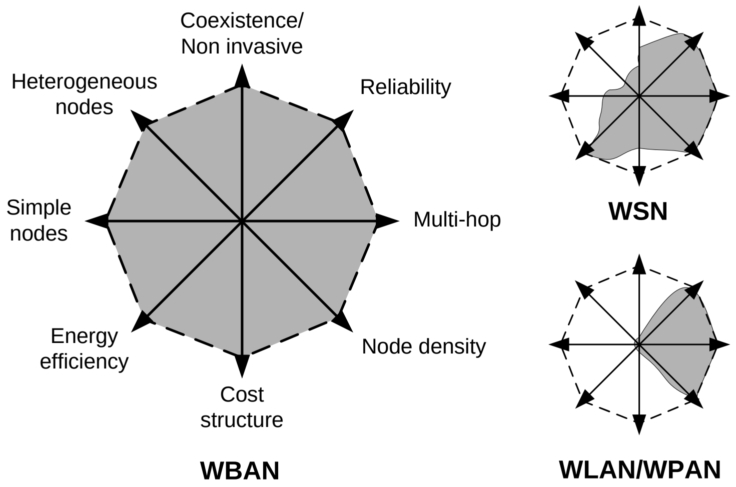


Figure 1.3: Characteristics of a Wireless Body Area Network compared with Wireless Sensor Networks (WSN) and Wireless Local Area Network (WLAN). Based on [33]

- The devices are located on the human body that can be in motion. WBANs should therefore be robust against frequent changes in the network topology;
- The propagation of the waves takes place in or on a (very) lossy medium like the human body. As a result, the waves are attenuated considerably before they reach the receiver. A relative simple but accurate propagation model is required;
- The data often consists of medical information. Hence, high reliability and low delay is required.
- Stringent security mechanisms are required in order to ensure the strictly private and confidential character of the medical data;
- And finally the devices are often very heterogeneous and may have very different demands or may require different resources of the network in terms of data rates, power consumption and reliability.

Currently no network protocols that meet (most of) the aforementioned requirements exist. Although a lot of research is being done toward energy efficient routing in ad hoc networks and sensor networks, the proposed solutions are inadequate for WBANs, as maximal throughput and minimal routing overhead are considered to be more important than minimal energy consumption. Energy efficient ad-hoc network protocols only attempt to find routes in the network that minimize energy consumption in terminals with small energy resources, thereby neglecting parameters such as the amount of operations (measurements, data processing, access to memory) required and the energy needed to transmit and receive a useful bit over the wireless link. Most protocols for WSNs only consider networks with homogeneous sensors and a many-to-one communication paradigm. In many cases the network is considered as a static one. In contrast, a WBAN has heterogeneous mobile devices with stringent real-time requirements due to the sensor-actuator communication. The communication covers a much smaller area and a WBAN has fewer devices. In addition, the vicinity of the human body imposes harsh requirements on the transmitted power and propagation loss. A schematic overview of the challenges in a WBAN and its comparison with other network types such as WSNs and WLANS/WPANs is given in Figure 1.3.

To summarize, the main challenges of a WBAN can be categorized in tackling the energy consumption and taking into account the reliability, the heterogeneity and the mobility of the network.

1.3 Main Research Contributions and Outline

From the problems and challenges mentioned in the previous section, we will mainly consider the energy consumption of the network, the reliability of the communication and the heterogeneity of the devices.

In a first step, the physical layer of the communication near the human body was studied and investigated. A path loss model that considers both line of sight communication and non-line of sight communication (i.e. from the front to the back of the body) is presented. This work was mainly done by Günter Vermeeren, Laurens Roelens and Elisabeth Reusens of the WiCa Research Group at Ghent University. In this work, this path loss model was used to investigate the connectivity between two nodes on the body. Instead of assuming a circular coverage area, a lognormal distribution is used to account for the signal fluctuations caused by irregularities in the surroundings of the receiving and transmitting antennas (aka shadowing effects). A probability function that models the connectivity is presented, giving a more realistic view of the network topology. Further, an energy model for the radio in WBANs is proposed. This energy model considers the receive and transmit energy of the radio. The transmitter energy contains 2 parts: a static part and a part for the transmitter amplifier that scales with the distance. Based on the previously proposed path loss model, the parameters for 3 frequently used radios are determined.

The connectivity model and energy model for the radio have been used to investigate possible network topologies for a WBAN. Two important characteristics were considered: energy efficiency and reliability. The study of minimizing the energy consumption (i.e. maximizing the lifetime of the network) considers both the lowering of the transmit power by using a multi-hop topology and by using aggregation of the packets toward the personal device. Next, the nodes cooperate in delivering the data to the personal device. This is analyzed via an ILP (integer linear programming) formulation. Relay devices are introduced to lower the maximum network energy even further. The influence of reliability on the network topology has been evaluated by setting up an ILP.

Depending on the specific application of the WBAN (i.e. medical monitoring, sport monitoring, gaming etc.) some features such as power control, data aggregation and so on will be required, others not. For each layer, an application developer has to determine which protocols are most suited for the purposes of the intended application, and has to go through the complicated process of combining them into an optimized protocol stack. Therefore, we present MOFBAN, a lightweight framework for a network architecture so as to allow for an easy and flexible adaptation of network protocols for a WBAN. The different components (the modules) are discussed. For communication between nodes, we present the modular header structure.

Once the network architecture is given, we study the network protocols for WBANs. In particular, the maximum throughput and minimum delay of IEEE 802.15.4 are determined and we discuss its suitability for WBANs. Further we present a new cross layer protocol for wireless Body Area Networks that both handles channel medium access and routing. For this purpose, a spanning tree is set up in a distributed manner and timeslots are used. Every node sends out a proprietary WASP-scheme to inform the nodes of the following level when they are allowed to send.

In order to counter these shortcomings, we propose CICADA. This protocol also sets up a spanning tree but defines a downwards and upwards cycle allowing bidirectional communication and lower delay. Reliability improvements and security mechanisms are added.

The development of WASP and CICADA has been done in close collaboration with Bart Braem of the University of Antwerp in the scope of the FWO-BAN[1], and the IBBT-IM3[2] projects.

The thesis is structured as follows. The taxonomy and the special requirements of a WBAN are briefly discussed in Chapter 2. The physical layer is investigated and the connectivity model and energy model of the radio are defined. These are used in Chapter 3 where the best network topology for a WBAN is identified using an ILP. The modular framework MOFBAN is presented in Chapter 4. In Chapter 5 the MAC-protocols for WBANs are analyzed, including the IEEE 802.15.4 protocol. After that, a first network protocol is presented, WASP. An improved version is presented in Chapter 6, called CICADA. The final Chapter 7 summarizes our scientific contributions and briefly discusses interesting future work.

[1] the Fund for Scientific Research - Flanders (F.W.O.-V.,Belgium) project G.0531.05

[2] Interactive Medical Mobile Monitoring, https://projects.ibbt.be/im3

1.4 Publications

The research results obtained during this PhD research have been published in scientific journals and presented at a series of international conferences. The following list provides an overview of the publications during my PhD research.

1.4.1 Publications in International Journals

[1] **B. Latré**, P. De Mil, I. Moerman, N. Van Dierdonck, B. Dhoedt, P. Demeester, *Maximum Throughput and Minimum Delay in IEEE 802.15.4*, published in Lecture Notes in Computer Science 3794, Proceedings of the 1st International Conference on Mobile Ad-hoc and Sensor Networks, MSN 2005, edited by X. Jia, J. Wu, Y. He, Vol. LNCS 3794:866–876, Wuhan, China, 13-15, December 2005

[2] **B. Latré**, P. De Mil, I. Moerman, N. Van Dierdonck, B. Dhoedt, P. Demeester, *Throughput and Delay Analysis of Unslotted IEEE 802.15.4*, Journal of Networks, ISSN 1796-2056, 1(1):20–28, May 2006

[3] E. Depoorter, **B. Latré**, I. Moerman, P. Demeester, *Symbiotic Networks: Towards a new level of cooperation in wireless networks*, published in Special Issue of the Wireless Personal Communications Journal, 45(4):479-495, Springer Publishers , June 2008

[4] **B. Latré**, B. Braem, C. Blondia, I. Moerman and P. Demeester, *Energy efficient and reliable networking in Wireless Body Area Networks*, submitted to Electronics Letters

[5] E. Reusens, W. Joseph, G. Vermeeren, L. Martens, B. Braem, C. Blondiam **B. Latré** and I. Moerman *Characterization of On-Body Communication Channel and Cross-Layer Design Application for Wireless Body Area Networks*, submitted to IEEE Transaction on Wireless Communications

1.4.2 Chapters in Books

[1] E. De Poorter, **B. Latré**, I. Moerman and P. Demeester, *Universal Framework for Sensor Networks*, Chapter in the book *Sensor and Ad-Hoc Networks: Theoretical and Algorithmic Aspects*, Series: Lecture Notes Electrical Engineering, Vol. 7, edited by Makki S. Kami, X.-Y. Li, et al. ISBN: 978-0-387-77319-3, In Press, Springer 2008 (available June 20)

1.4.3 Publications in International Conferences

[1] **B. Latré**, J. Hoebeke, L. Peters, T.Van Leeuwen, I. Moerman, B. Dhoedt and P. Demeester, *A heterogeneity based clustering heuristic for mobile ad hoc networks*, The 2004 IEEE International Conference on Communications (ICC 2004), pp. 3728–3733, Paris, France, June 20-24, 2004

[2] **B. Latré**, G. Vermeeren, I. Moerman, L. Martens, F. Louagie, S. Donnay and P. Demeester, *Networking and Propagation Issues in Body Area Networks*, 11th Symposium on Communications and Vehicular Technology in the Benelux (IEEE Benelux Chapter on Communications and Vehicular Technology), SCVT 2004, November 9, 2004, Ghent, Belgium

[3] W.Vandenberghe, **B. Latré**, F. De Greve, P. De Mil, S. Van den Berghe, K. Lamont, I. Moerman, M. Mertens, J. Avonts, C. Blondia and G. Impens, *A System Architecture for Wireless Building Automation*, published in Proceedings (on CD-ROM) of the 15th IST Mobile & Wireless Communications Summit 2006, Myconos, Greece, 4-8 June 2006

[4] B. Braem, **B. Latré**, C. Blondia, I. Moerman and P. Demeester, *The Wireless Autonomous Spanning Tree Protocol for Multihop Wireless Body Area Networks*, First International Workshop on Personalized Networks, San Jose, CA, 21 July 2006

[5] **B. Latré**, E. De Poorter, I. Moerman and P. Demeester, *Modular Architecture for Heterogeneous Sensor Networks*, 4th European conference on Wireless Sensor Networks (EWSN 2007), EWSN adjunct poster proceedings, pp. 21-22, Delft, Netherlands, 29-31 January 2007

[6] E. Reusens, W. Joseph, G. Vermeeren, L. Martens, **B. Latré**, B. Braem, C. Blondia, I. Moerman, *Path loss models for wireless communication channel along arm and torso: measurements and simulations*, published in Proceedings (on CD-ROM) of APS2007, the IEEE Antennas and Propagation Society International Symposium, pp. 345–348, Hawaii, USA, 10-15 June 2007,

[7] E. De Poorter, **B. Latré**, I. Moerman and P. Demeester, *Universal Framework for Sensor Networks*, International Workshop on Theoretical and Algorithmic Aspects of Sensor and Ad-hoc Networks (WTASA'07), Miami, Florida, June 28-29, 2007

[8] **B. Latré**, B. Braem, C. Blondia, I. Moerman and P. Demeester, *A Low-delay Protocol for Multi-hop Wireless Body Area Networks*, Proceedings (on CD-ROM) of the second International Workshop on Personal Networks (PerNets 2007), Philadelphia, USA, August 2007

[9] **B. Latré**, E. De Poorter, I. Moerman, P. Demeester, *Symbiotic networks*, published in Proceedings of the 9th Strategic Workshop on International Mobile Telecommunications, Malaga, Spain, 30 May - 1 June 2007

[10] B. Braem, **B. Latré**, C. Blondia, I. Moerman and P. Demeester, *The Need for Cooperation and Relaying in Short-Range High Path Loss Sensor Networks*, Proceedings of the IEEE International Conference on Sensor Technologies and Applications (Sensorcomm 2007), pp. 566–571, Valencia, Spain, October 2007

[11] **B. Latré**, E. De Poorter, I. Moerman and P. Demeester, *MOFBAN: a Lightweight Framework for Body Area Networks*, published in Lecture Notes in Computer Science 4808, Embedded and Ubiquitous Computing (EUC 2007), LNCS 4808:610–622, Taipei, Taiwan, December 2007

[12] W.Vandenberghe, **B. Latré**, F. De Greve, P. De Mil, S. Van den Berghe, K. Lamont, I. Moerman, M. Mertens, J. Avonts, C. Blondia and G. Impens, *A System Architecture for Wireless Building Automation*, published in KEIO and Gent University G-COE Joint workshop for future network, pp. 59–63, Ghent, Belgium, 20-21 March 2008,

[13] B. Braem, **B. Latré**, C. Blondia, I. Moerman and P. Demeester, *Improving Reliability in Multi-hop Body Sensor Networks*, accepted for presentation at the Second International Conference on Sensor Technologies and Applications (SENSORCOMM08), France, August 2008

[14] D. Singelée, **B. Latré**, B. Braem, M. De Soete, P. De Cleyn, B. Preneel, I. Moerman and C. Blondia, *A Secure Cross-layer Protocol for Multi hop Wireless Body Area Networks*, accepted for presentation at the 7th International Conference on AD-HOC Networks & Wireless (ADHOCNOW 2008), France, September 2008

1.4.4 Publications in National Conferences

[1] J. Hoebeke, **B. Latré**, I. Moerman, B. Dhoedt, P. Demeester, ”Routing in mobile Ad Hoc networks”, published in 4th FTW PHD Symposium, Interactive poster session, paper nr. 30 (proceedings available on CD-Rom), Gent, Belgium, December 3, 2003.

[2] **B. Latré**, I. Moerman, B. Dhoedt and P. Demeester, *Networking in Body Area Networks*, published in 5th FTW PHD Symposium, Interactive poster session, paper nr. 30 (proceedings available on CD-Rom), 1st of December 2004

References

[1] D. Cypher, N. Chevrollier, N. Montavont, and N. Golmie. *Prevailing over wires in healthcare environments: benefits and challenges*. IEEE Communications Magazine, 44(4):56–63, April 2006.

[2] R. S. H. Istepanian, E. Jovanov, and Y. T. Zhang. *Guest Editorial Introduction to the Special Section on M-Health: Beyond Seamless Mobility and Global Wireless Health-Care Connectivity*. Information Technology in Biomedicine, IEEE Transactions on, 8(4):405–414, December 2004.

[3] C. C. Y. Poon, Yuan-Ting Zhang, and Shu-Di Bao. *A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-health*. IEEE Communications Magazine, 44(4):73–81, April 2006.

[4] K Van Dam, S. Pitchers, and M. Barnard. *Body Area Networks: Towards a Wearable Future*. In Proceedings of WWRF kick off meeting, Munich, Germany, 6-7 March 2001.

[5] R. Schmidt, T. Norgall, J. Mörsdorf, J. Bernhard, and T. von der G"un. *Body Area Network BAN–a key infrastructure element for patient-centered medical applications.* Biomedizinische Technik. Biomedical engineering, 47(1):365–368, 2002.

[6] I. Chlamtac, M. Conti, and J.. Liu. *Mobile ad hoc networking: imperatives and challenges.* Ad Hoc Networks, 1(1):13–64, 2003.

[7] B. Gyselinckx, C. Van Hoof, J. Ryckaert, R. F. Yazicioglu, P. Fiorini, and V. Leonov. *Human++: autonomous wireless sensors for body area networks*. In Custom Integrated Circuits Conference, 2005. Proceedings of the IEEE 2005, pages 13–19, September 2005.

[8] C. Otto, A. Milenkovic, C. Sanders, and E. Jovanov. *System Architecture of a Wireless Body Area Sensor Network for Ubiquitous Health Monitoring*. Journal of Mobile Multimedia, 1(4):307–326, 2006.

[9] B. Lo and Guang-Zhong Yang. *Body Sensor Networks: Infrastructure for Life Science Sensing Research.* In Life Science Systems and Applications Workshop, 2006. IEEE/NLM, pages 1–2, Bethesda, MD,, July 2006.

[10] S. Park and S. Jayaraman. *Enhancing the quality of life through wearable technology*. IEEE Engineering in Medicine and Biology Magazine, 22(3):41–48, May/June 2003.

[11] IEEE 802.15 WPAN Task Group 6 Body Area Networks.

[12] Elliot Krames. *Implantable devices for pain control: spinal cord stimulation and intrathecal therapies*. Best Practice & Research Clinical Anaesthesiology, 16(4):619–649, December 2002.

[13] International Diabetes Federation (IDF) [Online] http://www.idf.org/.

[14] World Health Organization [online] http://www.who.int.

[15] Huan-Bang Li, Ken-ichi Takizawa, Bin Zhen, and Ryuji Kohno. *Body Area Network and Its Standardization at IEEE 802.15.MBAN*. In Mobile and Wireless Communications Summit, 2007. 16th IST, pages 1–5, Budapest, Hungary,, July 2007.

[16] L. Theogarajan, J. Wyatt, J. Rizzo, B. Drohan, M. Markova, S. Kelly, G. Swider, M. Raj, D. Shire, M. Gingerich, J. Lowenstein, and B. Yomtov. *Minimally Invasive Retinal Prosthesis*. In Solid-State Circuits, 2006 IEEE International Conference Digest of Technical Papers, pages 99–108, February 2006.

[17] V. Shankar, A. Natarajan, S. K. S. Gupta, and L. Schwiebert. *Energy-efficient protocols for wireless communication in biosensornetworks*. In Personal, Indoor and Mobile Radio Communications, 2001 12th IEEE International Symposium on, volume 1, San Diego, CA, USA, September 2001.

[18] B. Latré, G. Vermeeren, I. Moerman, L. Martens, and P. Demeester. *Networking and Propagation Issues in Body Area Networks*. In 11th Symposium on Communications and Vehicular Technology in the Benelux, SCVT 2004, November 2004.

[19] E. Jovanov, D. Raskovic, A. O. Lords, P. Cox, R. Adhami, and F. Andrasik. *Synchronized physiological monitoring using a distributed wireless intelligent sensor system*. In Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE, volume 2, pages 1368–1371, September 2003.

[20] Stefan Drude. *Requirements and Application Scenarios for Body Area Networks*. In Mobile and Wireless Communications Summit, 2007. 16th IST, pages 1–5, Budapest, Hungary, July 2007.

[21] S. Brebels, S. Sanders, C. Winters, T. Webers, K. Vaesen, G. Carchon, B. Gyselinckx, and W. De Raedt. *3D SoP integration of a BAN sensor node*. In 2005. Proceedings. 55th Electronic Components and Technology Conference, pages 1602–1606, May/June 2005.

[22] B. Gyselinckx, R. Vullers, C. V. Hoof, J. Ryckaert, R. F. Yazicioglu, P. Fiorini, and V. Leonov. *Human++: Emerging Technology for Body Area Networks*. In Very Large Scale Integration, 2006 IFIP International Conference on, pages 175–180, October 2006.

[23] O. C. Omeni, O. Eljamaly, and A. J. Burdett. *Energy Efficient Medium Access Protocol for Wireless Medical Body Area Sensor Networks*. In Medical Devices and Biosensors, 2007. ISSS-MDBS 2007. 4th IEEE/EMBS International Summer School and Symposium on, pages 29–32, Cambridge, UK,, August 2007.

[24] R.W. Hoyt, J. Reifman, T.S. Coster, and M.J. Buller. *Combat medical informatics: present and future*. In Proceedings of the AMIA 2002 annual symposium, pages 335–339, San Antonio, TX, November 2002.

[25] E. Jovanov, A. Milenkovic, C. Otto, and P. C. de Groen. *A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation*. Journal of NeuroEngineering and Rehabilitation, 2(1):16–23, March 2005.

[26] Moteiv [online] http://www.moteiv.com.

[27] T. Gao, D. Greenspan, M. Welsh, R. R. Juang, and A. Alm. *Vital Signs Monitoring and Patient Tracking Over a Wireless Network*. In Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the, pages 102–105, Shanghai,, 2005.

[28] K. Lorincz, D. J. Malan, T. R. F. Fulford-Jones, A. Nawoj, A. Clavel, V. Shnayder, G. Mainland, M. Welsh, and S. Moulton. *Sensor Networks for Emergency Response: Challenges and Opportunities*. IEEE Pervasive Computing, 3(4):16–23, 2004.

[29] K. Venkatasubramanian, G. Deng, T. Mukherjee, J. Quintero, V. Annamalai, and S.K.S. Gupta. *Ayushman: A Wireless Sensor Network Based Health Monitoring Infrastructure and Testbed*. In Distributed Computing in Sensor Systems, volume Volume 3560/2005, pages 406–407. Springer Berlin / Heidelberg, 2005.

[30] A. T. van Halteren, R. G. A. Bults, K. E. Wac, D. Konstantas, I. A. Widya, N. T. Dokovski, G. T. Koprinkov, V. M. Jones, and R. Herzog. *Mobile Patient Monitoring: The Mobihealth System*. The Journal on Information Technology in Healthcare, 2(5):365–373, October 2004.

[31] T. Falck, J. Espina, J. P. Ebert, and D. Dietterle. *BASUMA - the sixth sense for chronically ill patients*. In Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on, April 2006.

[32] IBBT IM3-project [online] http://projects.ibbt.be/im3.

[33] T. Zasowski. *A System Concept for Ultra Wideband (UWB) Body Area Networks*. PhD thesis, PhD Thesis, ETH Zürich, No. 17259, 2007. The thesis can be order directly via the publisher's webpage: http://www.logos-verlag.de/cgi-local/buch?isbn=1715.

2 Research Context

In this section, a general overview of existing research in the field of wireless Body Area Networks is given. We start with determining the taxonomy and requirements and discussing the different applications for a WBAN. Further, we point out the differences between a wireless sensor network and a WBAN. The general principles of networking in a WBAN are discussed. Next, the physical layer is studied: the propagation and link connectivity is considered and a radio model for WBANs is proposed.

2.1 Taxonomy and Requirements

From the applications described in Chapter 1, it is clear that a WBAN consists of several heterogeneous devices. This section will give an overview of the different types of devices used in a WBAN and discusses the requirements such as energy consumption, data rates and so on.

2.1.1 Different Types of Devices

We start with some definitions to capture the terminology:

(Wireless) Sensor node: a device that responds to and gathers data on physical stimuli, computes the data if necessary and reports this information wirelessly. It is made up out of several components: sensor hardware, a power unit, a processor, memory and a transmitter or transceiver [1].

(Wireless) Actuator node: a device that acts according to data received from the sensors or through interaction with the user. The components of an actuator are similar to the sensor's: actuator hardware (e.g. hardware for medicine administration, including a reservoir to hold the medicine), a power unit, a processor, memory and a receiver or transceiver.

(Wireless) Personal Device (PD): a device that gathers all the information acquired by the sensors and actuators and informs the user (i.e. the patient, a nurse, a GP etc.) via an external gateway, an actuator or a display/LEDS on the device. The components are a power unit, a (large) processor, memory and a transceiver. This device is also called a Body Control Unit (BCU) [2], body-gateway or a sink. In some implementations, a Personal Digital Assistant (PDA) or smartphone is used.

(Wireless) Relay Device: a device that acts like a router in the network, no actual sensing is done locally. The use of these devices will be further explained. The device holds the following components: a power unit, a (large) processor, memory and a transceiver.

Many different types of sensors and actuators are used in a WBAN. The main use of all these devices is to be found in the area of health applications, see Section 1.1.1. In the following, the term *nodes* refers to both the sensor as actuator nodes. The number of nodes in a WBAN is limited by nature of the network. It is expected that the number of nodes will be in the range of 20–50 [3, 4].

2.1.2 Data Rates

Due to the strong heterogeneity of the applications this will vary strongly, ranging from simple data at a few kbit/s to video streams of several Mbit/s. If possible, the data could be sent in bursts, which means that the data is sent at greater rate during the bursts.

In order to grasp the realm of data rates in a WBAN, we will consider some health care applications in a WBAN such as an electrocardiogram (ECG) that measures the activity of the heart or an electromyogram (EMG) that measures the activity of a muscle. The data rate of these applications can be calculated by means of the sampling rate, the range and the desired accuracy of the measurements [5, 6]. The sampling rate is defined as twice the required bandwidth according to the Shannon sampling theorem [9]. This gives us:

$$\text{Data rate} \; = \; \text{nr of bits} \cdot 2 \cdot f_{max} \qquad (2.1)$$

The data rates for the different applications are given in in Table 2.1. Overall, the application data rates are not high. However, if one has a WBAN with several of these devices (i.e. a dozen motion sensors, ECG, EMG, glucose monitoring etc.)

Application	Data Rate	Bandwidth	Accuracy	Reliability
ECG (12 leads, HF)	288 kbps	100-1000 Hz	12 bits	10^{-10}
ECG (6 leads, LF)	71 kbps	100-500 Hz	12 bits	10^{-10}
EMG	320 kbps	0-10,000 Hz	16 bits	10^{-10}
EEG (12 leads)	43.2 kbps	0-150 Hz	12 bits	10^{-10}
Blood saturation	16 bps	0-1 Hz	8 bits	10^{-10}
Glucose monitoring	1600 bps	0-50 Hz	16 bits	10^{-10}
Temperature	120 bps	0-1 Hz	8 bits	10^{-10}
Motion sensor	35 kbps	0-500 Hz	12 bits	10^{-3}
Cochlear implant	100 kbps	–	–	10^{-5}
Artificial retina	50-700 kbps	–	–	10^{-5}
Audio	1 Mbps	–	–	10^{-5}
Voice	50-100 kbps	–	–	10^{-3}

Table 2.1: Examples of medical WBAN applications [5–8]

the aggregated data rate easily exceeds 1 Mbps, which is a higher than the raw bit rate of most low power radios.

The reliability of the data transmission is provided in terms of the necessary bit error rate (BER) which is used as a= measure of number of lost packets. For a medical device, the reliability depends on the data rate. Low data rate devices can cope with a high BER (e.g. 10^{-4}), while devices with a higher data rate require a lower BER (e.g. 10^{-10}). The required BER is also dependent on the criticalness of the data. However it is not clear how the required BER should be determined exactly. The BER values in Table 2.1 should therefore be considered as guide numbers.

Two different types of data traffic are distinguished: *data control* and *regular data* traffic. The regular data traffic only represents the measurements made by the sensors sent to the central device. The data control traffic on the other hand covers all the other traffic, such as (battery) status information from the different sensors and actuators sent to the central device, tasks sent from the central device to the actuators, request for data from the central device to a sensor, settings from a sensor to an actuator [1],. . . With data control traffic we thus do not understand the overhead generated by the various protocols; this is referred to as (network) control traffic. The most important traffic pattern is the transport of the measured

[1] An example of a sensor sending data to an actuator is a sensor node that measures the glucose level in the blood. The sensor processes the data and if necessary the sensor sends a signal to an actuator that injects insulin

data between sensor and central device. Sensors in a WBAN can be considered as generating CBR (constant bit rate) traffic where traffic requirements do not change quickly over time [5].

The user or actuator node is not always interested in receiving data continuously from all devices at all times, but may only have interest for some specific events or measurements. As a consequence, four types of data delivery models can be defined [10]: *event-driven data delivery* where communication between the sensor and another device is only set up in case something happens or a threshold is reached. For example, if a sensor node notices the blood pressure crossing some threshold, a signal is sent to an actuator which can administer medication. *Demand-driven data delivery* only sends data when asked for. The third type sends his data continuously or periodically (i.e. every 10 seconds, every minute, . . .) to the interested devices. The time of sending the data is thus predetermined and is called *continuous-data rate* applications. The last type is *hybrid data delivery* which forms a combination of the three above. These different types of data delivery should be accounted for in the development of a routing protocol for a WBAN as each type requires his own approach. Most applications will have continuous data transmission. To be able to set the thresholds in the sensor nodes or to request information, it is obvious that communication from the personal device to the sensor node should be possible. This communication can be seen as control data and requires a low bandwidth.

2.1.3 Energy

The size of the battery used to store the needed energy is in most cases the largest contributor to the sensor device in terms of both dimensions and weight. Batteries are, as a consequence, kept small and energy consumption of the devices needs to be reduced. In some applications, a WBAN's sensor/actuator node should operate while supporting a battery life time of months or even years without intervention. For example, a pacemaker or a glucose monitor would require a lifetime lasting more than 5 years. Especially for implanted devices, the lifetime is very crucial. The need for replacement or recharging induces a cost and convenience penalty which is undesirable not only for implanted devices, but also for larger ones. The lifetime of a node for a given battery capacity can be enhanced by scavenging energy during the operation of the system. If the scavenged energy is larger than the average consumed energy, such systems could run eternally. However, energy scavenging will only deliver small amounts of energy [11, 12]. A combination of lower energy consumption and energy scavenging is the optimal solution for achieving autonomous Body Area Networks. For a WBAN, energy scavenging from on-body sources such as body heat and body vibration seems very well suited. In the former, a thermo electronic generator (TEG) is used to transform the temperature difference between the environment and the human body into electrical energy [7]. The latter uses for example the human gait as energy source [13].

Energy consumption can be divided into three domains: sensing, (wireless) communication and data processing [1]. During these processes the devices produce heat which is absorbed by the surrounding tissue and increases the temperature of the body. In order to limit this temperature rise and in addition to save the battery resources, the energy consumption should be restricted to a minimum. The amount of power absorbed by the tissue is expressed by the specific absorption rate (SAR). Since the device may be in close proximity to, or inside, a human body, the localized SAR could be quite large. The localized SAR into the body must be minimized and needs to comply with international or local SAR regulations. The regulation for transmitting near the human body is similar to the one for mobile phones, with strict transmit power requirements [14, 15]

The wireless communication is likely to be the most power consuming. In order to lower the energy consumption of the communication, the power of the radio transceiver can be reduced and the distance between the sender and receiver can be lowered by introducing relay devices and switching to multi-hop instead of working single-hop. This is the topic of Chapter 3. Other possible techniques for power saving are [16, 17]:

- avoid unnecessary retransmissions;
- put receiver in standby mode or turn radio off (sleep mode) whenever possible to minimize idle listening or overhearing;
- reduce frequency of sending network control messages;
- optimize size of message headers;
- use power efficient error correction schemes;
- routing based on the heterogeneity of the devices, shifting the load and power consumption to more capable devices;
- work on a lower carrier frequency, which reduces the absorption and hence the power needed;
- cross layer design where information of the different OSI-protocol layers is exchanged amongst the layers. As the protocol stack in a WBAN should be very small in order to minimize the energy consumption, the cross layer design is an important approach (see Chapter 4).

2.1.4 Quality of Service and Reliability

A proper quality of service (QoS) handling is an important part in the framework of risk management of medical applications. The critical factor is the reliability of the transmission, meaning that it is crucial that messages with monitoring information are received correctly by the health care professionals. The reliability can be considered either end to end or on a per link base. Examples of reliability include the guaranteed delivery of data (i.e. packet delivery ratio), in-order-delivery, . . . Besides that, messages should be delivered in reasonable time. The reliability of the network directly affects the quality of patient monitoring and in a worst case scenario it can be fatal when a life threatening event has gone undetected [18].

Other QoS parameters such as minimum guaranteed bandwidth, low delay and support of real time communication have a strong impact on MAC and PHY layers.

The desired quality of service will affect the energy consumption. For example, to obtain a lower packet loss, the transmit power can be increased, which raises the energy consumption. It is therefore important to achieve the right balance between power consumption and the desired reliability of the system.

2.1.5 Usability

In most cases, a WBAN will be set up in a hospital by medical staff, not by ICT-engineers. Consequently, the network should be capable of configuring and maintaining itself automatically, i.e. self-organization an self-maintenance should be supported. Whenever a node is put on the body and turned on, it should be able to join the network and set up routes without any external intervention. The self-organizing aspect also includes the problem of addressing the nodes. They can use a preconfigured address given at manufacturing time (e.g. the MAC-address) or an address given at setup time by the network itself. Further, the network should be quickly reconfigurable, i.e. for adding new services. When a route fails, a back up path should be set up (self-healing).

The nodes should have a small form factor consistent with wearable and implanted applications. This will make WBANs invisible and unobtrusive.

2.1.6 Computing

To be able to execute their duties, the devices will need some sort of embedded processor. The available computing power depends on the amount of energy present and process capabilities. Further, a device needs memory capacity. This is limited by the size of the device and the available energy. Consequently, only little data can be stored, which will lead to smaller interburst intervals. One has also to bear in mind that the protocol stack of the networking protocol requires a certain amount of memory. Thus, to limit the energy consumption, the protocol stack has to be sufficiently simple.

2.1.7 Place and Mobility

The devices may be scattered over and in the whole body. The exact location of a device will depend on the application, e.g. a heart sensor obviously must be placed in the neighborhood of the heart, a temperature sensor can be placed almost anywhere. Researchers seem to disagree on the ideal body location for some sensor nodes, i.e. motion sensors, as the interpretation of the measured data is not always the same [19]. The network should not be regarded as a static one. The body may be in motion (e.g. walking, running, twisting etc.) which induces channel fading and shadowing effects.

2.1.8 Security and Privacy

The communication of health related information between sensors in a WBAN and over the Internet to servers is strictly private and confidential [20] and should be encrypted to protect the patient's privacy. Furthermore, the medical staff who collects the data need to be confident that the data is not tampered with and indeed originates from that patient. It can not be expected that an average person or the medical staff plays the role of a network administrator who can set up and manage authentication and authorization processes. Moreover the network should be accessible when the user is not capable of giving the password (e.g. accessibility in trauma situations by the paramedics).

In [21] an algorithm based on biometric data is described, that can be employed to ensure the authenticity, confidentiality and integrity of the data transmission between the personal device and all the other nodes. Biometric is a technique commonly known as the automatic identification and verification of an individual by his or her physiological characteristics. Another method is presented in [22] where body-coupled communication (BCC) is used to associate new sensors in a WBAN. A BCC-transceiver generates a weak electric field that is capacitively coupled to the body. This allows transmitting small amounts of information. BCC can now be used by a new sensor to identify the person it belongs to, to discover all other sensors attached to the same person and to exchange network parameters.

Security and privacy protection mechanisms use a significant part of the available energy and should therefor be energy efficient and lightweight. Although providing sufficient security is a crucial factor in the acceptance of WBANs, we will not focus on the security mechanisms. This dissertation is limited to the networking issues and will only briefly address possible security measures.

2.1.9 Conclusion

From the analysis above, it is clear that we have a set of different requirements for supporting communication in a WBAN. In order to ease the development of

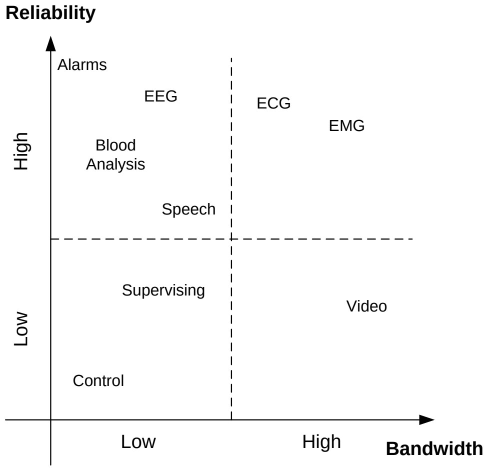


Figure 2.1: Traffic analysis in a WBAN.

new and robust applications, new protocols and a framework that supports these protocols are needed. These should consider the energy efficiency, the reliability, the different QoS-levels (i.e. delay), the heterogeneity (i.e. data rates, complexity), the ease of use of the system and the security/privacy requirements. In this dissertation, we mainly focus on the networking issues, and more specifically the energy efficiency and the reliability.

Figure 2.1 shows a visualization of different applications in terms of data rate and reliability, based on the information of Table 2.1. The applications can be grouped into 4 categories:

- Low data rate and low reliability
- Low data rate and high reliability
- High data rate and low reliability
- High data rate and high reliability

It is clear that the traffic in a WBAN is quite heterogeneous and newly developed protocols should be capable of coping with this heterogeneity in mind.

2.2 Networking in Body Area Networks

2.2.1 Wireless Sensor Networks

Body Area Networks can be considered as a special type of a wireless sensor networks (WSN) or a wireless sensor and actuator network (WSAN) with its own re-

Table 2.2: Schematic overview of differences between Wireless Sensor Networks and Wireless Body Area Networks [23].

Challenges	Wireless Sensor Network	Wireless Body Area Network
Scale	As large as the environment being monitored (meters / kilometers)	As large as the human body (meter)
Node Number	Greater number of nodes required for accurate, wide area coverage	Fewer, more accurate sensors node required (limited by space)
Node Function	Multiple sensors, each performing a dedicated task	Single sensor performing multiple tasks
Node Accuracy	Large node number compensates for accuracy and allows result validation	Limited node number, each required to be robust and accurate
Node Size	Small size preferable but not a major limitation in many cases	Need for small form factor
Event Detection	Early adverse event detection desirable, failure often reversible	Early adverse event detection vital, human failure irreversible
Node mobility	Much more likely to have a fixed or static structure	More variable topology due to body movement
Data Protection	Lower level wireless data transfer security required	High level wireless data transfer security required to protect patient information
Power Supply	Accessible and likely to be changed more easily and frequently	Inaccessible and difficult to replace in an implantable setting
Power Demand	Likely to be greater as power is more easily supplied	Likely to be lower as energy is more difficult to supply
Energy scavenging	Solar and wind power are the most likely candidates	Motion (vibration) and thermal (body heat) most likely
Access	Sensor more easily replaceable or even disposable	Implantable sensor replacement difficult
Biocompatibility	Not a consideration in most applications	A must for implants and some external sensors
Context Awareness	Not so important with static sensors where environments are well defined	Very important because body physiology is very sensitive to context change
Wireless Technology	Bluetooth, ZigBee, GPRS, WLAN, . . .	Low power wireless required
Data transfer	Loss of data during wireless transfer is likely to be compensated by redundant number of sensors used	Loss of data is more significant, and may require additional measures to ensure QoS and real-time data delivery. No redundant sensors available.

quirements[2]. Traditional sensor networks do not tackle the specific challenges associated with human body monitoring. The human body consists of a complicated internal environment that responds to and interacts with its external surroundings, but is in a way *separate* and *self-contained*. The human body environment has not only a smaller scale, but also requires a different type and frequency of monitoring, with different challenges than those faced by WSNs. The monitoring of medical data results in an increased demand for reliability. The ease of use of the sensors placed on the body leads to a small form factor that includes the battery and antenna part, resulting in a higher need for energy efficiency. Sensor nodes can move with regard to each other, for example a sensor node placed on the wrist in relation to a sensor node attached to the hip. This requires mobility support. In brief, although challenges faced by WBANs are in many ways similar to WSNs, there are intrinsic differences between the two, requiring special attention. An overview of some of these differences is given in Table 2.2.

2.2.2 Communication in Body Area Networks

The development and research in the domain of Body Area Networks is only at an early stage. As a consequence, the terminology is not always clearly defined. In literature, protocols developed for Body Area Networks can span from communication between the sensors on the body to communication from a body node to a data center connected to the Internet. In order to have clear understanding, we propose the following definitions: *intra-body communication* and *extra-body communication*. An example is shown on Figure 2.2. The former controls the information handling on the body between the sensors or actuators and the personal device [24–27], the latter ensures communication between the personal device and an external network [19, 28–30]. Doing so, the medical data from the patient at home can be consulted by a physician or stored in a medical database. This segmentation is similar to the one defined in [30] where a multi-tiered telemedicine system is presented. Tier 1 encompasses the intra-body communication, tier 2 the extra-body communication between the personal device and the Internet and tier 3 represents the extra-body communication from the Internet to the medical server. The combination of intra-body and extra-body communication can be seen as an enabler for ubiquitous health care service provisioning. An example can be found in [31] where Utility Grid Computing is combined with a WBAN to provide access to appropriate computational services and high bandwidth to a large collection of distributed time-varying resources.

To date, development has been mainly focused on building the system architecture and service platform for extra-body communication. Much of these imple-

[2] In the following, we will not make a distinction between a WSAN and a WSN although they have significant differences.

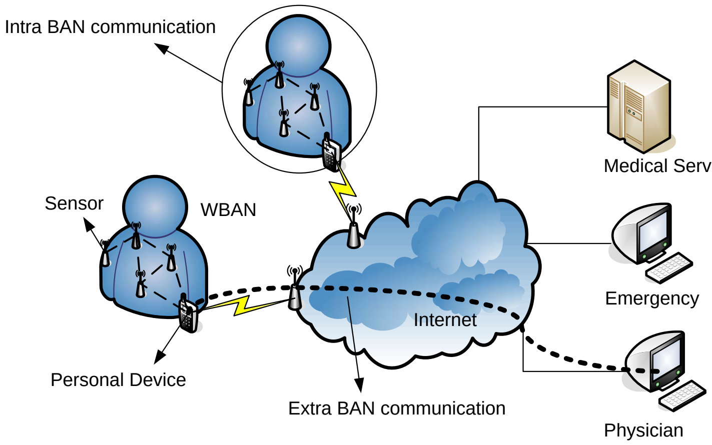


Figure 2.2: Example of intra-body and extra-body communication in a WBAN

mentations focus on the repackaging of traditional sensors (e.g. ECG, heart rate) with existing wireless devices. They consider a very limited WBAN including only a few sensors directly and wirelessly connected to a personal device. Further they use transceivers with a large form factor and large antenna that are not adapted for use on a body. Hence, the routing aspects of these protocols are very minimal as only direct communication is to be contemplated.

The focus of this thesis is intra-body communication. Developing efficient routing protocols in WBANs is a nontrivial task because of the specific characteristics of a wireless environment. First of all, the available bandwidth is limited, shared and can vary due to fading, noise and interference, so the protocol's amount of network control information should be limited. Secondly, the nodes that form the network can be very heterogeneous in terms of available energy or computing power. A discussion of existing routing protocols will be given in Section 5.2. Tests with TelosB-motes (using the CC2420 transceiver) showed lack of communications between nodes located on the chest and nodes located on the back of the patient. This was accentuated when the transmit power was set to a minimum for energy savings reasons. Similar conclusions with a CC2420 transceiver where drawn in [32]: when a person was sitting on a sofa, no communication was possible between the chest and the ankle. Better results were obtained when the antenna was placed 1 cm above the body. As the devices get smaller and more ubiquitous, a direct connection to the personal device will no longer be possible and more complex network topologies will be needed. In the following section, we will discuss the characteristics of the propagation in a WBAN.

2.3 Physical Layer

In this section, the physical layer in Body Area Networks will be studied. First, the different types of propagation will be discussed and a simple path loss model will be given. Based on this model, the probability of connectivity between two radios is discussed. Finally, a first order radio model is constructed.

2.3.1 Propagation in Body Area Networks

Several researchers have been investigating the path loss along and inside the human body either using narrowband radio signals or Ultra Wide Band (UWB). All of them came to the conclusion that the radio signals experience great losses. Generally in wireless networks, it is known that the transmitted power drops off with d^η where d represents the distance between the sender and the receiver and η the coefficient of the path loss (aka propagation coefficient) [33, 34]. In free space, η has a value of 2. Other kind of losses include fading of signals due to multipath propagation. The propagation can be classified in propagation inside the body and along the body. In this thesis, we focus on the latter.

2.3.1.1 Propagation in the Body

The propagation of electromagnetic (EM) waves in the human body has been investigated in [35, 36]. The body acts as a communication channel where losses are mainly due to absorption of power in the tissue, which is dissipated as heat. As the tissue is lossy and mostly consists of water, the EM-waves are attenuated considerably before they reach the receiver. In order to determine the amount of power lost due to heat dissipation, they use the specific absorption rate (SAR), a standard measure of how much power is absorbed in the tissue. It is concluded that the path loss is very high and that, compared to the free space propagation, an additional 30-35 dB at small distances is noticed. A simplified temperature increase prediction scheme based on SAR is presented in [36]. It is argued that considering energy consumption is not enough and that the tissue is sensitive to temperature increase. The influence of a patient's body shape and position on the radiation pattern from an implanted radio transmitter has been studied in [37]. It is concluded that the difference between different body shapes (male, female and child) are at least as large as the impact of a patient's arm movements.

Next to the propagation of radio waves, several researchers have investigated the possibility to transfer electronic data by capacitive and galvanic coupling, also called body-coupled communication (BCC). These radios work at low frequencies (ranging from 10 kHz to 10 MHz). Zimmerman [38] first showed the potential of interference-free ultra low power data communication through the human body. High variations of the transmission attenuation have been observed at different locations of the body. Galvanic coupling promises to be a potential communication technology for sensor application on the thorax and for short distances on the limbs [39]. This technology can also be used to exchange data from one body to another by for example shaking hands [40]. In [41] OsteoConduct is presented, where the human musculoskeletal system is used to transmit data and information in a low-power, secure, non-intrusive fashion. Although this research looks promising, only very small data rates can be achieved (5 bits/second).

The idea of BCC is further exploited by [22] for bootstrapping wireless Body Area Networks. They argue to equip the nodes with both RF and BCC capabilities. As a BCC is restricted to a person's body, the BCC can be used to discover and identify sensor nodes on the same body and for waking up RF radios from low-power sleep mode.

2.3.1.2 Propagation along the Body

Most of the devices used in a WBAN however are attached on the body. The propagation along the human body can be divided into line of sight (LOS) and non-line of sight (NLOS) situations. In the former, the curvature effects of the body are not taken into account as simulations are performed on a flat phantom or experiments are done at one side of the body. In the latter, the effect of propagation from the front of the body to the side or back are evaluated.

The channel model for line of sight (LOS) propagation along the human body was studied in [42–45], both by simulations and experiments. The studies were done for either narrowband and UWB signals, but the results are similar. It was found that the path loss exponent η is between 3 and 4, depending on the position of the device, i.e. the path loss on the arm is lower than the one on the trunk. They claim that it is probably due to the higher absorption in the larger volume of the trunk, and because the surface of the trunk is less flat than the surface of the stretched arm. The study in [44] shows a significant impact of the antenna height on the path loss. The closer the antenna is to the body, the higher the path loss: a difference of more than 20 dB is found for an antenna placed at 5 mm and 5 cm. As the sensors and antennas of a Body Area Network will be designed to be as small as possible, the antenna will be close to the body which will result in a higher path loss.

In non-line of sight (NLOS) situations, there is no direct view between the sender and receiver. The EM-waves are diffracting around the body rather than having a direct path through the body. In [45, 46], a path loss exponent ranging from 5 to 6 was found and [47] shows the same behavior for the antenna position as in LOS situations. We observe a higher path loss along the NLOS channel than along the LOS channel, due to diffraction around the human body[3] and absorp-

[3] This is also referred to as creeping waves. A creeping wave is defined as the wave that is diffracted

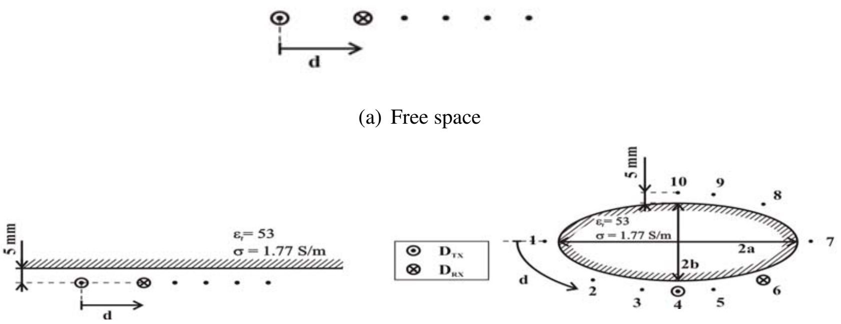
tion of a larger amount of radiation by the body. In [48] the dominant propagation mechanism for the the ear-to-ear link, which can be regarded as a worst case scenario at the head due to the missing line-of-sight component, was identified. It was shown that direct transmission can be neglected for transmission from one side of the head to the opposite due to reflections and the strong attenuation of the head. Further, the movement of the body plays an important role. In [49] a preliminary system model for gait analysis has been proposed. It is concluded that significant attenuation can occur (up to 20 dB) when a body limb is moved in between the Tx and Rx antenna. According to [50] the movement of the limbs can induce an attenuation of 30 dB or more. A similar conclusion was found in an actual implementation [27] where the sensors communicate directly with the personal device using an RF-radio operating at 868 MHz. Loss rates of more than 50% where found when the body was in motion.

Study of Propagation: configuration As an example, we examine the following scenario where the propagation loss of a path between two dipoles in the proximity of a conducting dielectric body is studied[4]. Two half-wavelength dipoles radiating at 2400 MHz were placed at a distance of 5 mm in front of a flat phantom (or conducting dielectric body), and around an elliptical cylindrical phantom representing a torso of a human body through which we could investigate the curvature effects. The former represents the LOS situation and the latter the NLOS path loss. Figure 2.3 shows these two configurations together with a reference configuration: two aligned dipoles in free space. The height, the gap and radius of the dipole are 56.25 mm, 1 mm and 1 mm, respectively. The flat and elliptical cylindrical phantom have the dielectric properties of average muscle as found in the FCC properties database [51]: ϵ_r = 53 and σ = 1.77 S/m. The length of the major and minor axes of the elliptical cross-section measures respectively 350 mm and 260 mm. In the free-space and the flat-phantom configurations, the transmitting dipole stays at a fixed position and the receiving dipole shifts away up to a distance of 500 mm. In the elliptical cylindrical configuration, the transmitting dipole is placed at the positions $i = 1$ to 4 (Figure 2.3(c)). The receiving dipole is located at $j = i + 1$ to $i + 6$ for each position of the transmitting dipole. The configurations are simulated with the commercial electromagnetic FDTD solver SEMCAD.

Study of Propagation: discussion Figure 2.4 plots the variation of the ratio of the received power (P_r) and the transmitted power (P_s), at the terminals of the dipole antennas, as a function of the distance between the transmitting and

around the shadowed surface of a smooth body such as a sphere.

[4] This study was done in collaboration with Günter Vermeeren from the INTEC-WiCa research group. The results were published in [46].



(a) Free space

(b) Flat phantom (LOS)

(c) Cylindrical phantom (NLOS)

Figure 2.3: Propagating paths between dipoles [46]

receiving dipole antenna. A simple approximating model for this ratio is given by:

$$\frac{P_r}{P_s} = G_{rec}G_{send}(\frac{\lambda}{4\pi d})^{\eta} \tag{2.2}$$

with G_{rec} and G_{send} the gain of the used receiving and sending antenna, λ the free-space wavelength, d the distance between the dipoles and η the path loss coefficient. For $\eta = 2$, (2.2) transforms into the well-known Friis transmission formulation [33] for free-space propagation. The ratios of the powers extracted from the simulations are fitted to (2.2). The results are shown on Figure 2.4. For free-space propagation (reference configuration), the fit returns a path loss coefficient η of 2, as expected. For the case of the flat phantom, η increases already to 4.6. A worst-case loss coefficient of 5.8 is obtained when the dipoles are radiating in the proximity of the elliptical cylindrical phantom. We also notice the variation of the path loss for different positions of the transmitting dipole. Due to the differences in curvature the propagated waves experience a different attenuation between transmitting and receiving dipole. When the receiving dipole reaches position $j = i+4$ to $i+6$, transmitting and receiving dipoles are not in Line-Of-Sight (LOS). As a consequence, the path loss increases at a higher rate than in the case of the flat phantom.

2.3.1.3 Path Loss Model

This preliminary study has shown that the path loss around the human body thus tremendously exceeds the path loss for propagation in free space and that the path loss differs for LOS and NLOS communication. Due to these high losses, direct communication between the sensors and the sink will not always be possible, especially when one wants to lower the radio's transmission power.

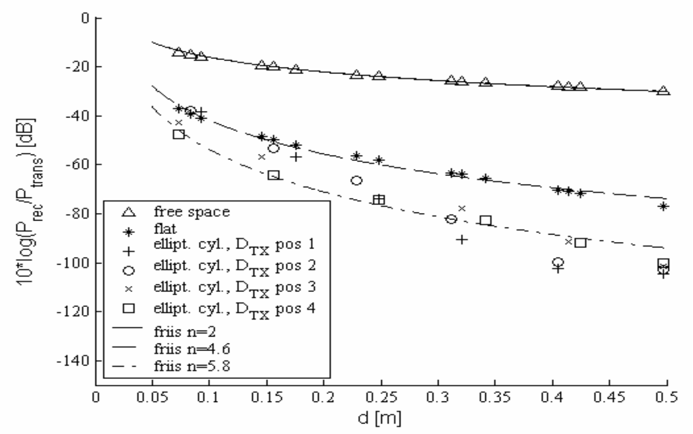


Figure 2.4: The ratio of received power and transmitted power as a function of distance [46]

parameter	value LOS [42]	value NLOS [47]
d_0	10 cm	10 cm
$P_{0,dB}$	35.7 dB	48.8 dB
σ	6.2 dB	5.0 dB
η	3.38	5.9

Table 2.3: Parameter values for the path loss model (2.3)

To model the propagation between the transmitting and the receiving antenna as a function of the distance d, we use the following semi-empirical formula for the path loss, as in [42, 47]:

$$PL_{dB}(d) \; = \; PL_{0,dB} \; + \; 10 \cdot \eta \cdot \log_{10}(\frac{d}{d_0}) \tag{2.3}$$

where $PL_{0,dB}$ is the path loss at a reference distance d_0 and η is the path loss exponent. The parameter values are found by fitting the formula to the measurements obtained from simulations. Table 2.3 gives an overview of the fitted results for two different propagation channels and the variation σ of the individual measurements around the model. The first channel is located along the front of the torso and is LOS [42]. The second channel is measured around the torso, resulting in NLOS propagation [47]. We ignore the influence of the movement of the limbs or other parts of the body as the studies in that area are still too premature and not accurate.

The model in (2.3) only represents the mean path loss [34]. However, in practice the average received power varies from location to location in an apparently random manner. The magnitude of the standard deviation indicates how strong the signal fluctuates caused by irregularities in the surroundings of the receiving and transmitting antennas. This variation is well described by a lognormal distribution with standard deviation σ and is called *shadowing* [33]. It is crucial to account for this in order to provide a certain reliability of communications. The total path loss then becomes a random variable given by

$$PL_{dB}(d) \;=\; PL_{0,dB} \;+\; 10 \cdot \eta \cdot \log_{10}(\frac{d}{d_0}) + X_{\sigma,dB} \tag{2.4}$$

where $X_{\sigma,dB}$ represents the shadowing component and is a zero-mean Gaussian random variable in dB with standard deviation σ. When σ = 0, we notice that (2.4) reduces to the normal path loss model of (2.3). The different values of σ for the path loss models can be found in Table 2.3.

2.3.2 Modeling Link Connectivity

In order to have correct reception of a radio signal, it is necessary that the received power of the signal is more than a threshold value $P_{th,dB}$. This value depends on the parameters of the radio and the background noise. Correct reception means that all the bits are correctly received. The received signal strength $P^j_{r,dB}$ at a node j from a node i sending at $P^i_{s,dB}$ with a distance d_{ij} can be found using (2.4):

$$P^j_{r,dB}(d_{ij}) \;=\; P^i_{s,dB} \;-\; PL_{dB}(d_{ij}) \tag{2.5}$$

The second term represents a random variable, which means that $P^j_{r,dB}$ will also be a random variable. In [52] and [53] a probabilistic model for connectivity based on a lognormal model of power variations has been proposed. Based on this study, we can define the following probability for link connectivity between two radios i and j:

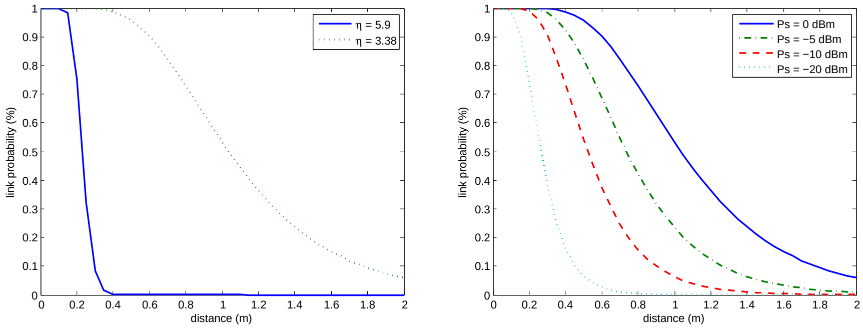
$$\begin{aligned} p(d_{ij}) &= \Pr\left[P^j_{r,dB}(d_{ij}) \;>\; P_{th,dB}\right] && (2.6)\\ &= \Pr\left[P^i_{s,dB} - PL_{0,dB} - 10\eta\log_{10}(\frac{d_{ij}}{d_0}) - X_{\sigma,dB} > P_{th,dB}\right] && (2.7)\\ &= \Pr\left[A_{dB}(d_{ij}) \;<\; 0\right] && (2.8) \end{aligned}$$

where

$$A_{dB}(d_{ij}) \;=\; X_{\sigma,dB} - P^i_{s,dB} + PL_{0,dB} + 10\eta\log_{10}(\frac{d_{ij}}{d_0}) + P_{th,dB} \tag{2.9}$$

In section 2.3.1.3 we have seen that the path loss is modeled as normally distributed. Consequently, (2.9) can be seen as normally distributed with standard deviation σ around the mean $\mu(d_{ij})$ where:

$$\mu(d_{ij}) \;=\; -P^i_{s,dB} + PL_{0,dB} + 10\eta\log_{10}(d_{ij}/d_0) + P_{th,dB} \tag{2.10}$$



(a) Varying η with $P_{s,dB}$ = 0 dBm

(b) Varying send power $P_{s,dB}$ with η = 3.38

Figure 2.5: Link probability of communication around the human body. For σ and d_0 we have used the values of Table 2.3, depending on the value of η.

Consequently, (2.6) can be rewritten as

$$p(d_{ij}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{0} \exp\left[-\frac{(t - \mu(d_{ij}))^2}{2\sigma^2}\right] dt \tag{2.11}$$

$$= \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{\mu(d_{ij})}{\sqrt{2}\sigma}\right) \tag{2.12}$$

The receiver threshold $P_{th,dB}$ is found by using the estimates for noise at the receiver. In many low power radios, the receiver sensitivity is about -90 dBm [54, 55]. If we require a signal-to-noise ratio (SNR) of at least 20 dB, we can set the receiver threshold to

$$P_{th,dB} = P_{sense} + SNR = -90 + 20 = -70\,\text{dBm} \tag{2.13}$$

The link probability for the path loss models in Table 2.3 is given in Figure 2.5. In Figure 2.5 (a) $P_{s,dB}$ is set to 0 dBm and the differences between LOS and NLOS are given. It can be seen that the link probability for the NLOS scenario is a lot lower with respect to the one for LOS for the same distance. This is expected due to the high losses in the NLOS scenario. We can clearly see that, due to the high attenuation near the body, the distance where the link probability is higher than 50% is very small. Figure 2.5 (b) shows the link probabilities for varying $P_{s,dB}$ when the LOS scenario is used. As expected, the probability lowers when a lower send power is used.

The probability of the packet loss can be modeled as the opposite of the link probability.

In the case that σ = 0, we notice that (2.12) reduces to

$$p(d_{ij}) = \begin{cases} 1, & \text{for } \mu(d_{ij}) < 0 \\ 0, & \text{for } \mu(d_{ij}) > 0 \end{cases} \tag{2.14}$$

Using (2.10) and the equation above, we find that in order to have connectivity, the following constraint must be met:

$$P^{i}_{s,dB} - PL_{0,dB} - 10\eta \log_{10}(d_{ij}/d_0) > P_{th,dB} \tag{2.15}$$

This is the regular connectivity model used in wireless networks where the connectivity between two nodes is only determined by their mutual distance [56]. Only when the nodes are within a threshold distance d_t of each other, they are connected: this is represented as a circular coverage area; d_t can be determined by using (2.15).

The link connectivity can be regarded as the probability that a packet is correctly received. Therefore, it can be used as an alternative for determining the packet error rate (PER) of the link.

2.3.3 Energy Model of the Radio

The results obtained in the previous section can be used to investigate the energy consumption of protocols and to determine the appropriate strategy when developing such a protocol. An important element for this is to dispose of a good energy model of the radio. In this section, we start with a short overview of general energy models used in wireless networks as currently no energy models for WBANs exist.

2.3.3.1 Existing Models

A much used first order energy model for radios in sensor networks was proposed by Heinzelman et al. [57]. A d^2 energy loss due to channel transmission is assumed with d the distance between sender and receiver. In order to transmit k bits over a distance d, the radio expends

$$E_{tx}(k,d) = E_{TXelec} \cdot k + E_{amp} \cdot k \cdot d^2 \tag{2.16}$$

$$E_{rx}(k) = E_{RXelec} \cdot k \tag{2.17}$$

In these formulas, E_{tx} represents the transmission energy (in J) for a packet of k bits over a distance d, E_{rx} the receiver energy. E_{TXelec} and E_{RXelec} are distance independent terms that account for the real world overhead of transmitter and receiver electronics respectively for one bit (PLLs, VCOs, biascurrents, etc.) and the digital processing. The second term in (2.16) describes the energy needed to send k bits over a distance d. E_{amp} is the energy consumed by the transmit amplifier. This is shown in Figure 2.6. The specific values of these parameters are hardware dependent and pertain to the sort of radio interface used. It is assumed that the radios have power control and consume the minimal energy needed to reach the receiver. This is also the main drawback as power control is not as simple as it seems.

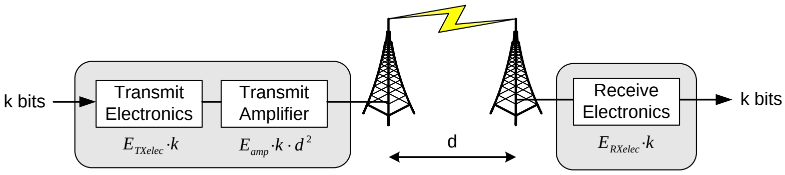


Figure 2.6: First order energy model for radios in wireless sensor networks [57].

A more extensive model is found in [58]: next to (2.16) and (2.17), 2 additional formulas are introduced covering the sensor core and the computation core:

$$E_{sense} = e_{sense} \cdot k \tag{2.18}$$

$$E_{comp} = n_{agg} \cdot e_{comp} \cdot k \tag{2.19}$$

where

e_{sense}	=	Energy to sense a bit
e_{comp}	=	Energy to compute a bit
n_{agg}	=	Number of aggregated streams.

In [59] a model is presented where the node decrements the available energy according to the following parameters: (a) the specific network interface controller characteristics, (b) size of the packets and (c) the bandwidth used. The following equations represent the energy used (in J) when a packet is transmitted (2.20) or received (2.21). The packet size is in bits.

$$E_{tx} = \frac{I_{tx} \cdot V}{Bandwidth} \cdot packetsize \tag{2.20}$$

$$E_{rx} = \frac{I_{rx} \cdot V}{Bandwidth} \cdot packetsize \tag{2.21}$$

where I_{tx} represents the current while transmitting and I_{rx} the current while transmitting. Although the equipment not only consumes energy while sending and receiving but also while listening, the models above assume that the listening operation is energy free. This model does not take a static energy consumption due to processing packets into account.

In [60], the energy models of existing network simulators such as GlomoSim and NS-2 are discussed. Further, a new energy model based on (2.20) and (2.21) is proposed. This model considers all possible radio operations, such as transmitting, receiving, overhearing, idle mode, sensing and sleeping. However, the proposed model is not given entirely.

2.3.3.2 The Model We Used

As energy model, we have chosen the first order model from (2.16) and (2.17). In section 2.3.1.3 we have seen that around the body the path loss exponent is higher

nRF2401 [54]		CC2420 [55]		RF 230 [61]	
RF tx power	DC tx current	RF tx power	DC tx current	RF tx power	DC tx current
0 dBm	13 mA	0 dBm	17.4 mA	3 dBm	15 mA
-5 dBm	10.5 mA	-1 dBm	16.5 mA	1 dBm	13 mA
-10 dBm	9.4 mA	-3 dBm	15.2 mA	-3 dBm	11 mA
-20 dBm	8.8 mA	-5 dBm	13.9 mA	-17 dBm	10 mA
		-7 dBm	12.8 mA		
		-10 dBm	11.2 mA		
		-15 dBm	9.9 mA		
		-20 dBm	8.5 mA		
DC rx current	Voltage	DC rx current	Voltage	DC rx current	Voltage
19 mA	1.9 V	18.8 mA	2.85 V	16 mA	3 V
Raw bit rate					
1 000 000 bps		250 000 bps		250 000 bps	

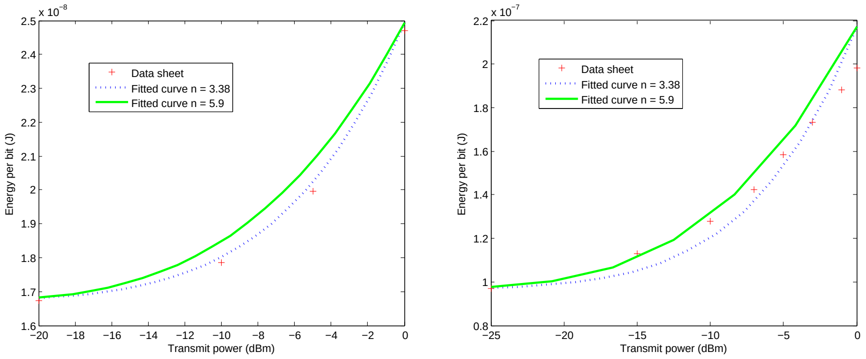
Table 2.4: Data sheet values for the Nordic nRF2401, the Chipcon CC2420 and ATMEL RF230 transceivers

than 2. Therefore, we have changed the formula to a more general one by replacing d^2 with d^η. Further, E_{amp} also varies with the loss coefficient, so $E_{amp}(\eta)$ is used instead. The specific values of these parameters are hardware dependent. We have determined these parameters for three commercially available transceivers which are frequently used in sensor networks: the Nordic nRF2401 low power single chip transceiver [54], the Chipcon CC2420 transceiver [55] used in a.o the Telos-B and MicaZ motes and the ATMEL RF230 transceiver [61]. All transceivers work in the 2.4–2.45 GHz band and have a very low power consumption.

As we are only interested in the energy consumption of the communication, which is much larger than the energy used for sensing [62], we ignore the latter in the following discussion.

Table 2.4 gives an overview of the sending and receiving characteristics of the transceivers. The receive energy E_{RXelec} can easily be obtained by looking at the current while receiving and the raw bit rate. For example, the Nordic transceiver has a receive current of 19 mA, which gives a power consumption of 36.1 mW or 36.1 nJ/bit. The power consumption of the transmitter E_{tx} can be calculated using the DC transmit current and the voltage level of the transceiver. In the following, we will determine the values of the parameters of (2.16) using the calculated values of E_{tx}.

For the distance in (2.16), we use the maximal distance that can be reached between the sender and the receiver. This means that if the receiver is positioned a little bit further, it can no longer hear the sender. This distance can be calculated for each output power level (P_{tx}) using (2.4) and the assumption that the maximal path loss (P_{dB}) equals the difference between the sensitivity of the radio and P_{tx}. However, the shadowing effect discussed in section 2.3.1.3 also plays an important



(a) Fitting for the nRF 2401

(b) Fitting for the CC 2420

Figure 2.7: The energy consumption of the nRF 2404 and the CC2420 transceivers.

parameter	nRF2401	CC2420	RF 230
E_{TXelec}	16.72 nJ/bit	96.9 nJ/bit	120 nJ/bit
E_{RXelec}	36.1 nJ/bit	172.8 nJ/bit	192 nJ/bit
$E_{amp}(3.38)$	0.2314e-9 J/bit	2.964e-7 J/bit	0.82e-9 J/bit
$E_{amp}(5.9)$	1.565e-6 J/bit	10.54e-4 J/bit	5.6026e-6 J/bit

Table 2.5: Parameter values for the Nordic nRF2401, the Chipcon CC2420 and ATMEL RF 230 transceivers

role as the maximum distance can vary over time. Taken this into account would mean that also the energy model would become probabilistic. For reasons of simplicity and in order to have a usable energy model, the shadowing effect is ignored for determining the parameter values of Table 2.5. Thus, $\sigma = 0$. So for each output level, we calculate the maximal distance. Next, we use this distance and the corresponding E_{tx} calculated from the data in Table 2.4 to fit (2.16). An example of the fitting can be seen in Figure 2.7. It can be seen that the first order model fits rather well the actual energy consumption of the transmitter. The results for the radios for different values of the path loss exponent can be found in Table 2.5. It can be seen that the Nordic radio has a significantly lower energy consumption per bit. This can be explained by the higher bitrate that can be obtained by the Nordic transceiver. Hence, we will use the parameters of the Nordic radios in our further calculations.

2.4 Conclusion

In this chapter, we have discussed the taxonomy and requirements of a WBAN. A WBAN imposes some specific challenges and requirements on the network, such as energy efficiency, reliability, heterogeneity etc. New protocols are needed to support these requirements. Further, we have pointed out the differences between a WSN and a WBAN. Although these types of networks seem very similar, intrinsic differences can be found. For example, in the scale of the network, even more stringent energy requirements, mobility support, heterogeneity, dynamics and so on.

The physical layer of a WBAN was discussed thoroughly. An overview of existing studies about the propagation in and along the human body was given and a path loss model was presented for LOS and NLOS communication. It was concluded that the electromagnetic waves along the body experience a high path loss. The path loss model is subsequently used to model the connectivity with a lognormal distribution that accounts for the shadowing effects. This connectivity model can be used as a measure to determine the reliability of the network. As a last item, an energy model for the radio has been proposed. This energy model considers the receive and transmit energy of the radio. The transmitter energy contains two parts: a static part and a part for the transmitter amplifier that scales with the distance. Based on the previously proposed path loss model, the parameters for three frequently used radios are determined.

References

[1] I. F. Akyildiz, Weilian Su, Y. Sankarasubramaniam, and E. Cayirci. *A survey on sensor networks*. IEEE Communications Magazine, 40(8):102–114, August 2002.

[2] R. Schmidt, T. Norgall, J. Mörsdorf, J. Bernhard, and T. von der G"un. *Body Area Network BAN–a key infrastructure element for patient-centered medical applications.* Biomedizinische Technik. Biomedical engineering, 47(1):365–368, 2002.

[3] T. Zasowski, F. Althaus, M. Stager, A. Wittneben, and G. Troster. *UWB for noninvasive wireless body area networks: channel measurements and results*. In Ultra Wideband Systems and Technologies, 2003 IEEE Conference on, pages 285–289, November 2003.

[4] C. Otto, A. Milenkovic, C. Sanders, and E. Jovanov. *System Architecture of a Wireless Body Area Sensor Network for Ubiquitous Health Monitoring*. Journal of Mobile Multimedia, 1(4):307–326, 2006.

[5] T. Penzel, B. Kemp, G. Klosch, A. Schlogl, J. Hasan, A. Varri, and I. Korhonen. *Acquisition of biomedical signals databases*. IEEE Engineering in Medicine and Biology Magazine, 20(3):25–32, May/June 2001.

[6] S. Arnon, D. Bhastekar, D. Kedar, and A. Tauber. *A comparative study of wireless communication network configurations for medical applications*. IEEE [see also IEEE Personal Communications] Wireless Communications, 10(1):56–61, February 2003.

[7] B. Gyselinckx, J. Penders, and R. Vullers. *Potential and challenges of body area networks for cardiac monitoring, Issue 6, Supplement 1, , ISCE 32nd Annual Conference, November-December 2007, Pages S165-S168.* Journal of Electrocardiolog, 40(6):S165–S168, November-December 2006. ISCE 32nd Annual Conference.

[8] L. Theogarajan, J. Wyatt, J. Rizzo, B. Drohan, M. Markova, S. Kelly, G. Swider, M. Raj, D. Shire, M. Gingerich, J. Lowenstein, and B. Yomtov. *Minimally Invasive Retinal Prosthesis*. In Solid-State Circuits, 2006 IEEE International Conference Digest of Technical Papers, pages 99–108, February 2006.

[9] C. E. Shannon. *Communication in the presence of noise*. Proceedings of the IEEE, 72(9):1192–1201, September 1984.

[10] S. Tilak, N. B. Abu-Ghazaleh, and W. Heinzelman. *A taxonomy of wireless micro-sensor network models*. SIGMOBILE Mob. Comput. Commun. Rev., 6(2):28–36, 2002.

[11] J. A. Paradiso and T. Starner. *Energy Scavenging for Mobile and Wireless Electronics*. IEEE Pervasive Computing, 04(1):18–27, 2005.

[12] B. Gyselinckx, C. Van Hoof, J. Ryckaert, R. F. Yazicioglu, P. Fiorini, and V. Leonov. *Human++: autonomous wireless sensors for body area networks*. In Custom Integrated Circuits Conference, 2005. Proceedings of the IEEE 2005, pages 13–19, September 2005.

[13] T. von Buren, P. D. Mitcheson, T. C. Green, E. M. Yeatman, A. S. Holmes, and G. Troster. *Optimization of inertial micropower Generators for human walking motion*. IEEE Sensors Journal, 6(1):28–38, February 2006.

[14] International Commission on Non-ionizing Radiation Protection (ICNIRP). *Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz)*. Health Physics, 74(4):494–522, apr 1998.

[15] *IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz*. 1999.

[16] C-K Toh, editor. *Ad Hoc Mobile Wireless Networks: Protocols and systems*. Prentice Hall PTR, Englewood Cliffs, NJ, 2002.

[17] I. Chlamtac, M. Conti, and J.. Liu. *Mobile ad hoc networking: imperatives and challenges*. Ad Hoc Networks, 1(1):13–64, 2003.

[18] U. Varshney and S. Sneha. *Patient monitoring using ad hoc wireless networks: reliability and power management*. IEEE Communications Magazine, 44(4):49–55, April 2006.

[19] E. Jovanov, A. Milenkovic, C. Otto, and P. C. de Groen. *A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation*. Journal of NeuroEngineering and Rehabilitation, 2(1):16–23, March 2005.

[20] A. Bhargava and M. Zoltowski. *Sensors and wireless communication for medical care*. In Database and Expert Systems Applications, 2003. Proceedings. 14th International Workshop on, pages 956–960, September 2003.

[21] C. C. Y. Poon, Yuan-Ting Zhang, and Shu-Di Bao. *A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-health*. IEEE Communications Magazine, 44(4):73–81, April 2006.

[22] T. Falck, H. Baldus, J. Espina, and K. Klabunde. *Plug 'n play simplicity for wireless medical body sensors*. Mob. Netw. Appl., 12(2-3):143–153, 2007.

[23] Guang-Zhong Yang, editor. *Body Sensor Networks*. Springer-Verlag London Limited, 2006.

[24] B. Latré, B.Braem, I.Moerman, C. Blondia, E. Reusens, W. Joseph, and P. Demeester. *A Low-Delay Protocol for Multihop Wireless Body Area Networks*. In Mobile and Ubiquitous Systems: Networking & Services, 2007 4th Annual International Conference on, Philadelphia, PA, USA, August 2007.

[25] T. Watteyne, S. Augé-Blum, M. Dohler, and D. Barthel. *AnyBody: a Self-organization Protocol for Body Area Networks*. In Second International Conference on Body Area Networks (BodyNets), Florence, Italy, 11-13 June 2007. 2007.

[26] D. Takahashi, Y. Xiao, and F. Hu. *LTRT: Least Total-Route Temperature Routing for Embedded Biomedical Sensor Networks*. In IEEE Globecom 2007, November 2007.

[27] A. Ylisaukko-oja, E. Vildjiounaite, and J. Mantyjarvi. *Five-Point Acceleration Sensing Wireless Body Area Network - Design and Practical Experiences*. iswc, 00:184–185, 2004.

[28] N. T. Dokovski, A. T. van Halteren, and I. A. Widya. *BANip: Enabling Remote Healthcare Monitoring with Body Area Networks*. In N. Guelfi, E. Astesiano, and G. Reggio, editors, FIDJI 2003 International Workshop on Scientific Engineering of Distributed Java Applications, Luxembourg, volume 2952/2004 of *Lecture notes in Computer Science*, pages 62–72, Heidelberg, 2004. Springer Verlag.

[29] K. E. Wac, R. Bults, A. van Halteren, D. Konstantas, and V. F. Nicola. *Measurements-based performance evaluation of 3G wireless networks supporting m-health services*. In S. Chandra and N. Venkatasubramanian, editors, Multimedia Computing and Networking 2005. Edited by Chandra, Surendar; Venkatasubramanian, Nalini. Proceedings of the SPIE, Volume 5680, pp. 176-187 (2004)., pages 176–187, December 2004.

[30] A. Milenkovic, C. Otto, and E. Jovanov. *Wireless sensor networks for personal health monitoring: Issues and an implementation*. Computer Communications, Wireless Sensor Networks and Wired/Wireless Internet Communications, 29(13-14):2521–2533, August 2006.

[31] O. O. Olugbara, M. O. Adigun, S. O. Ojo, and P. Mudali. *Utility Grid Computing and Body Area Network as Enabler for Ubiquitous Rural e-Healthcare*

Service Provisioning. In e-Health Networking, Application and Services, 2007 9th International Conference on, pages 202–207, Taipei, Taiwan,, June 2007.

[32] R. C. Shah and M. Yarvis. *Characteristics of on-body 802.15.4 networks*. In Wireless Mesh Networks, 2006. WiMesh 2006. 2nd IEEE Workshop on, pages 138–139, Reston, VA, USA,, 2006.

[33] T. S. Rappaport. *Wireless Communication: Principles and Practice 2nd edition*. Prentice Hall, 2002.

[34] S. R. Saunders. *Antennas and propagation for wireless communication systems*. Wiley, West Sussex, 1999.

[35] S. K. S. Gupta, S. Lalwani, Y. Prakash, E. Elsharawy, and L. Schwiebert. *Towards a propagation model for wireless biomedical applications*. In Communications, 2003. ICC '03. IEEE International Conference on, volume 3, pages 1993–1997, May 2003.

[36] Q Tang, N. Tummala, S. K. S. Gupta, and L. Schwiebert. *Communication scheduling to minimize thermal effects of implanted biosensor networks in homogeneous tissue*. IEEE Transactions on Biomedical Engineering, 52(7):1285–1294, July 2005.

[37] A. J. Johansson. *Wave-propagation from medical implants-influence of body shape on radiation pattern*. In [Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] EMBS/BMES Conference, 2002. Proceedings of the Second Joint, volume 2, pages 1409–1410, 2002.

[38] T. Zimmerman. *Personal area networks: Nearfield intrabody communication*. IBM Systems Journal, 35(3):609–617, 1996.

[39] M. S. Wegmueller, A. Kuhn, J. Froehlich, M. Oberle, N. Felber, and W. Kuster, N.and Fichtner. *An Attempt to Model the Human Body as a Communication Channel*. IEEE Transactions on Biomedical Engineering, 54(10):1851–1857, October 2007.

[40] K. Hachisuka, Y. Terauchi, Y. Kishi, T. Hirota, K. Sasaki, H. Hosaka, and K. Ito. *Simplified circuit modeling and fabrication of intrabody communication devices*. In Solid-State Sensors, Actuators and Microsystems, 2005. Digest of Technical Papers. TRANSDUCERS '05. The 13th International Conference on, volume 1, pages 461–464, June 2005.

[41] L. Zhong, D. El-Daye, B.t Kaufman, N. Tobaoda, T. Mohamed, and M. Liebschner. *OsteoConduct: Wireless body-area communication based on bone conduction*. In Proc. Int. Conf. Body Area Networks (BodyNets), June 2007.

[42] E. Reusens, W. Joseph, G. Vermeeren, and L. Martens. *On-body Measurements and Characterization of Wireless Communication Channel for Arm and Torso of Human*. In International Workshop on Wearable and Implantable Body Sensor Networks (BSN'07), pages 26–28, Aachen, March 2007.

[43] E. Reusens, W. Joseph, G. Vermeeren, L. Martens, B. Latré, B. Braem, C. Blondia, and I. Moerman. *Path-Loss Models for Wireless Communication Channel along Arm and Torso: Measurements and Simulations*. In IEEE AP-S Internation Symposium 2007, JUN 2007.

[44] L. Roelens, S. Van den Bulcke, W. Joseph, G. Vermeeren, and L. Martens. *Path loss model for wireless narrowband communication above flat phantom*. International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06), 42(1):10–11, January 2006.

[45] T. Zasowski, G. Meyer, F. Althaus, and A. Wittneben. *Propagation effects in UWB body area networks*. In Ultra-Wideband, 2005. ICU 2005. 2005 IEEE International Conference on, pages 16–21, September 2005.

[46] B. Latré, G. Vermeeren, I. Moerman, L. Martens, and P. Demeester. *Networking and Propagation Issues in Body Area Networks*. In 11th Symposium on Communications and Vehicular Technology in the Benelux, SCVT 2004, November 2004.

[47] A. Fort, J. Ryckaert, C. Desset, P. De Doncker, P. Wambacq, and L. Van Biesen. *Ultra-wideband channel model for communication around the human body*. IEEE Journal on Selected Areas in Communications, 24:927–933, April 2006.

[48] T. Zasowski. *A System Concept for Ultra Wideband (UWB) Body Area Networks*. PhD thesis, PhD Thesis, ETH Zürich, No. 17259, 2007. The thesis can be order directly via the publisher's webpage: http://www.logos-verlag.de/cgi-local/buch?isbn=1715.

[49] M. Di Renzo, R. M. Buehrer, and J. Torres. *Pulse Shape Distortion and Ranging Accuracy in UWBbased Body Area Networks for FullBody Motion Capture and Gait Analysis*. In IEEE Globecom 2007, November 2007.

[50] D. Neirynck. *Channel Characterisation and Physical Layer Analysis for Body and Personal Area Network Development*. PhD thesis, University of Bristol, UK, November 2006.

[51] FCC dielectric properties database, http://www.fcc.gov/fcc-bin/dielec.sh.

[52] R. Hekmat and P. Van Mieghem. *Connectivity in wireless ad-hoc networks with a log-normal radio model.* Mobile Networks and Applications, 11(3):351–360, 2006.

[53] C. Bettstetter and C. Hartmann. *Connectivity of wireless multihop networks in a shadow fading environment.* Wireless Networks, 11(5):571–579, 2005.

[54] Nordic, nRF 2401 datasheet [online] http://www.nordicsemi.com /index.cfm?obj=product&act=display&pro=64.

[55] Chipcon, CC2420 datasheet [online] http://focus.ti.com/docs/prod /folders/print/cc2420.html.

[56] C. Bettstetter. *On the minimum node degree and connectivity of a wireless multihop network.* In MobiHoc '02: Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing, pages 80–91, New York, NY, USA, 2002. ACM.

[57] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. *Energy-efficient communication protocol for wireless microsensor networks.* In System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on, January 2000.

[58] M. Bhardwaj and A. P. Chandrakasan. *Bounding the lifetime of sensor networks via optimal role assignments.* In INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, volume 3, pages 1587–1596, 2002.

[59] J. C. Cano and P. Manzoni. *A performance comparison of energy consumption for Mobile Ad HocNetwork routing protocols.* In Modeling, Analysis and Simulation of Computer and Telecommunication Systems, 2000. Proceedings. 8th International Symposium on, pages 57–64, San Francisco, CA, USA, 2000.

[60] C. B. Margi and K. Obraczka. *Instrumenting network simulators for evaluating energy consumption in power-aware ad-hoc network protocols.* In Modeling, Analysis, and Simulation of Computer and Telecommunications Systems, 2004. (MASCOTS 2004). Proceedings. The IEEE Computer Society's 12th Annual International Symposium on, pages 337–346, October 2004.

[61] ATMEL Corporation, RF230 datasheet [online] http://www.atmel.com/dyn/products/product_card.asp?part_id=3941.

[62] M. Welsh. *Exposing resource tradeoffs in region-based communication abstractions for sensor networks.* Computer Communication Review, 34(1):119–124, 2004.

3
Network Topologies for Wireless Body Area Networks

The network topology is the logical organization or arrangement of communication devices in the system. The selection of a proper network topology in wireless Body Area Networks is important as it significantly affects the overall system performance and protocol design. It influences the system in many ways such as power consumption, ability to handle heterogeneity, robustness against failures, the routing of data etc. This chapter will investigate the optimal network topology for WBANs where two design aspects will be considered: the energy efficiency and the reliability.

3.1 Introduction

Only few researchers have investigated the (optimal) network topology for WBANs. Most researchers assume that a single-hop topology where every sensor directly communicates with the personal device is the best solution. However, as pointed out in Section 2.2, the devices currently used are not fitted for a WBAN. When they get smaller and more ubiquitous, more complex network topologies will be adopted. This is supported by [1] where an energy comparison for a UWB-transmission from a node on the back to a node on the chest has shown that a multi hop strategy is recommended, if the energy for transmission and reception is higher than a ten thousandth part of the energy for bit encoding and decoding (i.e. the static energy). The difference between a cluster based approach and a tree based model was investigated in [2]. It was concluded that a cluster-based approach is the most energy efficient. However, they used a simplified propagation model with a fixed path loss of $\eta = 2$ and did not consider the effect of LOS and NLOS situations.

3.1.1 Energy Efficiency

As described in Section 2.1.3, energy consumption is an important issue in WBANs. Equipping the sensors with replaceable or rechargeable batteries is not always possible, as this reduces the comfort of the person wearing them. Further, in Section 2.3.1 we have seen that the human body is a very lossy medium and the electromagnetic waves are attenuated considerably. This also induces tissue heating [3, 4] which can arise when too much power is transmitted near the human body.

The energy efficiency can be measured by several performance indicators. In a WBAN, it is important that the network functions as long as possible. Further, all the nodes are equally important as no redundant nodes can be used (due to cost reasons and usability) and all the sensors generate and transport medical data. Therefore, we propose to use the network lifetime as primary metric, which we define as the time for the first node to die, i.e. to run out of energy. In order to have a high network lifetime, the most consuming node should be made more energy efficient. To that end, we define the *maximum network energy* as the energy consumed by the most consuming node(s) of the network. This metric is inversely proportional to the network lifetime. When the maximum network energy is high, the lifetime will be low and vice versa. The network lifetime metric forces us to consider all nodes to be equally important.

By decreasing the distance between transmitter and receiver, the energy consumption of a node can be lowered. In Figure 3.1 an example is shown for the energy consumption of the Nordic nRF2401 radio with the parameter of Table 2.4. The distance is calculated with (2.4) and a path loss η of 5.9. It is clear that by putting the nodes closer to each other, the energy consumption will drop. However, we can not keep on repeating this process as at a certain distance the energy gain will be very small, i.e. when the distance is less than 30 centimeter. Indeed, the static energy of the transceiver will consume more energy than the transmitter. The transmission distance between two nodes can be lowered by using intermediate nodes as relay devices. In order to know whether or not the use of relay devices is more energy efficient, the static energy needed for forwarding the data needs to be considered.

A study for a line topology in sensor networks has been done in [5]. They used a similar radio model and found that multi hop communication is better when

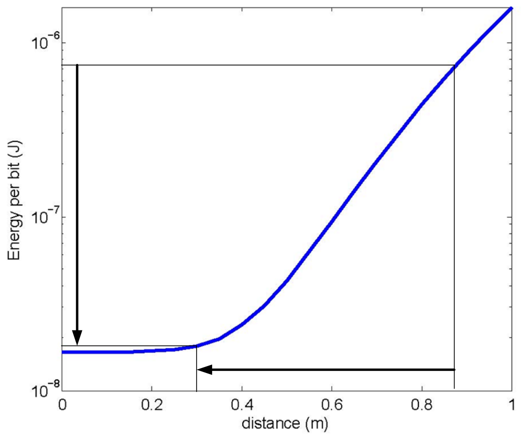


Figure 3.1: An example of reducing energy consumption in a WBAN. By decreasing the distance that needs to be covered, the output energy can be set lower and the overall energy consumption will drop. The distance is calculated with (2.4) and the path loss η = 5.9.

the distance between the sender and receiver exceeds the *characteristic distance* d_{char} [6]. Using the notations from this thesis, d_{char} is defined as

$$d_{char} = \sqrt[\eta]{\frac{E_{TXelec} + E_{RXelec}}{E_{amp} \cdot (\eta - 1)}} \tag{3.1}$$

For the parameters of Table 2.5 for the Nordic chip, we see that d_{char} is 43 cm for a η = 5.9. This distance is rather small compared to the size of the body, so it is expected that a multi hop architecture will be the most energy efficient. The optimal number of hops can be then be calculated as follows when the distance between the sender and receiver is D:

$$K_{opt} = \left\lceil \frac{D}{d_{char}} \right\rceil \tag{3.2}$$

In a WBAN, the maximum distance between two nodes will never by higher than 2 meter. Thus, the optimal number of hops will never be higher than 5 with d_{char} = 43 cm. This definition for d_{char} however only considers a line topology and does not take into account the effects of aggregation.

In this work, the energy efficiency will be studied in several stages. First, we investigate the effects of a multi hop approach analytically for two different topologies: a line and a tree topology. Then the most optimal network topology is determined using linear programming. Next, relay devices are added to the network and the ILP-formulation is extended. In a last analysis, cooperation between relay devices and normal devices is investigated.

3.1.2 Aggregation

While studying the energy efficiency, the effect of in-network aggregation will also be taken into account. Aggregation results in reducing the size of data that is forwarded by an intermediate node to its parent. This is an important topic for wireless networks in general. Data or packets are aggregated in order to reduce the number of packet transmissions. In wireless sensor networks, the measured data is highly correlated and aggregation is used to merge information. This merging function can be a very simple mathematical function, such as max, min or $average$, or it can be a very complex algorithm. This data merging is not usable in a WBAN as each sensor will measure a specific physiological parameter. Fasolo *et al.* give an overview of the aggregation techniques in WSNs [7]. Cluster based data-aggregation uses cluster heads which collect the data from surrounding neighbors. The cluster head performs local aggregation and sends the digest to the sink. A typical example is LEACH [8]. A second approach uses aggregation trees where data is aggregated while it is being routed over the aggregation path, e.g. Directed Diffusion [9].

In this work, we will consider aggregation where the payload from different packets is fused and a new common header is added. Each data packet has a header that contains the address of the source, the destination, error correction codes etc. When the node forwards the packet, this header can be omitted when the data is sent together with the data of the node itself as this data already contains a header. The length of the packet when it is forwarded is thus reduced by a factor ϕ. This is referred to as the *aggregation factor*. For example, if a node i receives a packet of k bits from node j and aggregates it with its own data (l bits), the new packet size of node i is $l + \phi k$ bits. When ϕ = 1, no aggregation takes place (e.g. when the packet is sent separately). This type of aggregation is also referred to as concatenation since the packets are concatenated in one packet.

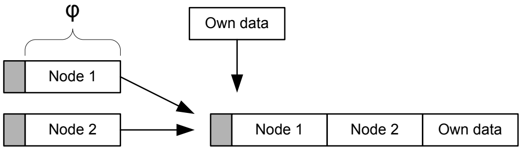


Figure 3.2: An example of aggregation in WBANs. The node receives data from nodes 1 and 2 with length 1. The header is removed, only the payload remains (length ϕ). This payload is combined with the node's own data to form a new packet and a new header is added. The length of the packet is now 2ϕ+1. If no aggregation is used, there would be three packets with length 1.

3.1.3 Reliability

The influence of the network architecture on the reliability in WBANs was briefly studied in [10]. The star and multi-hop topologies are evaluated against the packet delivery ratio in an experimental setting. The packet delivery ratio could by increased by either operating at a higher power level or by changing to a multi-hop architecture. As the energy consumption is restricted, using a multi-hop topology is the preferred way to go.

Reliability has also been studied in wireless sensor networks. The authors of [11] claim that using the shortest path metric is not enough and propose a routing metric based on the link quality in order to increase the reliability of the network. This metric is also used in [12] but has the tendency to generate long paths and it does not evaluate the energy consumption. A similar approach is used in [13] but tries to limit the number of retransmissions using a multi hop architecture. This leads to smaller and more energy efficient networks.

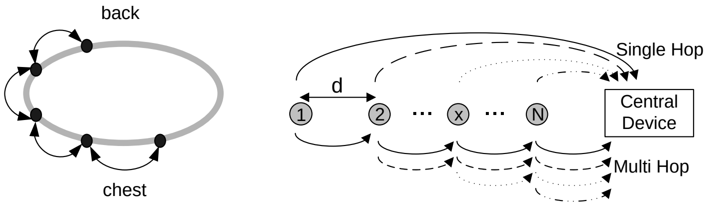
In this work, the reliability is studied for different metrics such as packet delivery ratio using the link probability and collision of probability. The influence of the energy efficiency on the reliability is investigated and the ILP-formulation is once again extended so that the most reliable and energy efficient path is chosen.

3.2 Multi-hop Network Versus a Star Network

In a multi-hop network the nodes are connected to the personal device (PD), possibly through other nodes. In a star network all nodes are directly connected to the sink or personal device. The PD takes all the necessary actions based on the collected data. In general, the nodes in a multi-hop network will transmit at lower power.

To study the effects of a multi-hop approach two different topologies are taken into account: in the first one all the nodes are equidistantly placed on one line (Line topology) and in the second one the nodes form a tree network (Tree topology). In both we assume that all nodes in the network generate packets at the same rate, so each duty cycle all nodes send one packet to the sink. In this study, a perfect duty cycle is assumed, i.e. a sensor only turns its radio on when it sends or receives data. The main purpose of this approach is to orthogonalize the results from this study and the properties of specific MAC-protocols.

Based on the propagation model described in section 2.4, we used a smaller path loss exponent for links between nearby nodes. This is done to take into account the curvature effects of the body as shown on Figure 3.3(a). The links between neighboring nodes can be modeled as a LOS situation. In other situations, the path loss for NLOS is used.



(a) Communication path on the body

(b) The resulting line topology

Figure 3.3: Communication paths on the body. Only the LOS paths are shown. All other paths are considered to be NLOS where a higher path loss exponent is used. The resulting line topology is shown on the right.

3.2.1 Line Topology

The first topology is very simple: all nodes are situated on one line with a fixed distance d between the nodes, as shown on Figure 3.3(b). Every node sends a packet of k bits to the sink. The counting starts at the node the farthest from the sink.

When there are N nodes in the network, the energy usage for sending a packet of k bits for node x in a star network (i.e. single hop) is the following, using the notations from Chapter 2:

$$E_{SH}(x,d) = E_{TXelec} \cdot k + E_{amp}(\eta) \cdot ((N - x + 1) \cdot d)^{\eta} \cdot k \tag{3.3}$$

The energy consumption when a multi-hop architecture is used, requires an extra term for the energy consumption of the receiver:

$$E_{MH}(x,d) = k \, [(x-1) \cdot E_{RXelec} + x \cdot (E_{TXelec} + E_{amp}(\eta) \cdot d^{\eta})] \tag{3.4}$$

Figure 3.4 shows the ratio of multi-hop energy usage over single-hop energy usage (E_{MH}/E_{SH}) for a scenario with 4 nodes and changing distances between the nodes. The parameters of the Nordic transceiver and different path losses, i.e. the LOS and NLOS values, are used. The results show that the nodes closest to the sink perform really bad when using multi-hop: they become *hotspots* using more than 10 times the energy of single-hop. This is because these nodes relay all the data from the lower nodes. However, far away from the sink, at node 1, single-hop performs up to 1000 times worse because of the high path loss. For the intermediate hops, the distance between the nodes and to the sink determines whether the single-hop or the multi-hop topology is the most energy efficient. It is therefore clear that distance plays an important role in these results. This is in line with the study performed in [5] where the use of multi-hop is more energy efficient when the distance between the node and the central device exceeds d_{char} (3.1), i.e. 43

cm in these scenarios. When the inter node distance increases, the difference between the LOS and NLOS propagation starts to impact performance dramatically.

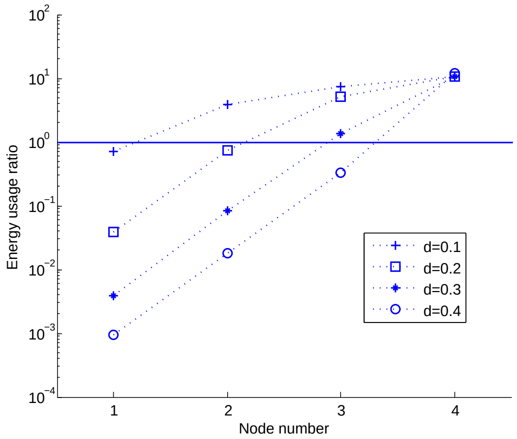


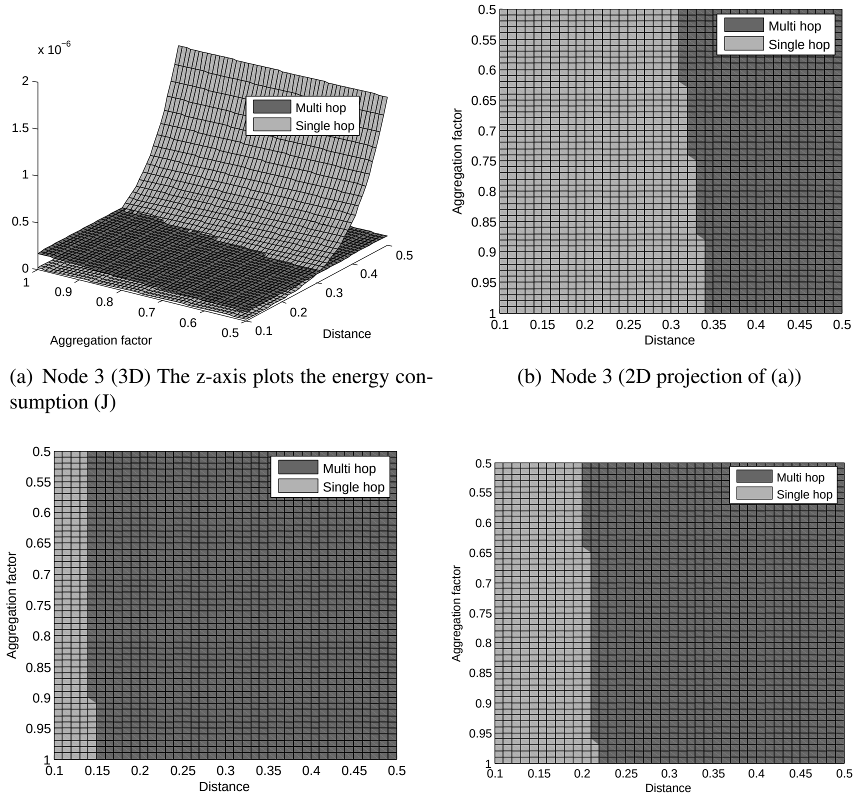
Figure 3.4: Comparison of the energy usage when using single-hop and multi-hop communication in the line scenario with 4 nodes of Fig. 3.3(b). The distance between the nodes (d) is varied from 0.1 to 0.4 m. The ratio of the energy usage of the multi-hop and single-hop communication is given (E_{MH}/E_{SH}). When the ratio is above the horizontal line, single-hop is the most energy efficient.

Besides the distance, aggregation of data plays an important role. When a node receives k bits, only a part of these bits are considered as *useful*, as was discussed in Section 3.1.2. Only ϕk bits (with ϕ the aggregation factor) of the k received bits will be forwarded to the next hop. The energy consumption for single-hop will remain the same. The energy consumption for multi-hop can be reformulated as:

$$E^{A}_{MH}(x,d) \;=\; k\cdot\sum_{i=0}^{x-2}\phi^{i}\cdot E_{RXelec} + k\cdot\sum_{i=0}^{x-1}\phi^{i}\cdot(E_{TXelec}+E_{amp}(\eta)\cdot d^{\eta}) \quad (3.5)$$

The summation can be proved by induction as follows. When node 2 receives k packets from node 1, only ϕk bits need to be forwarded. So node 2 will send $(1+\phi)k$ bits. In general, this can be written as $\sum_{i=0}^{x-1}\phi^{i}\cdot k$ for a node on position x. Assuming that this is true for x, we will proof the correctness for $x+1$. For a node on position $x+1$, the number of forwarded bits is then given by $[\sum_{i=0}^{x-1}\phi^{i}\cdot k]\cdot\phi+k$ or $k\cdot[\sum_{i=0}^{x-1}\phi^{i+1}+1]$ or $k\cdot[\sum_{i=0}^{x}\phi^{i}]$. As the formula holds for $x+1$ and for x = 2, the assumption was correct. If ϕ equals 1, (3.5) converts into (3.4).

Figure 3.5 shows the most optimal topology for a linear topology of 4 sensor nodes for nodes 1, 2 and 3 separately, for varying inter node distance and aggregation factor. Figure 3.5(a) plots the energy consumption for the multi-hop and single-hop scenario separately, for varying ϕ and d. The lowest surface has the



(a) Node 3 (3D) The z-axis plots the energy consumption (J)

(b) Node 3 (2D projection of (a))

(c) Node 1 (2D projection)

(d) Node 2 (2D projection)

Figure 3.5: Influence of data aggregation and inter node distance on the energy consumption for a linear topology of 4 nodes (Fig. 3.3(b)) for nodes 1, 2 and 3 separately. The lighter area means that a single-hop topology is more energy efficient. The distance is expressed in meters [m].

lowest energy consumption. This is shown in Figure 3.5(b), a 2D-projection of Figure 3.5(a). The lighter area means that a single-hop topology is more energy efficient. For nodes 1 and 2 only the 2D-projection is given. It can be seen that the influence of the aggregation factor is limited in a line topology.

3.2.2 Tree Topology

Now the nodes are arranged in a tree topology which is commonly used in sensor networks. Compared to a line topology, a tree based topology induces a larger forwarding overhead as each node has more children. In WBANs the formation of the tree and the number of children per level largely depends on the situation:

the number of nodes in the network, the position of the nodes,... In this analysis, we assume a ζ-balanced tree, i.e. a tree where all the nodes have ζ children. In Figure 3.6 we have depicted a balanced tree with ζ = 2, i.e. a binary tree. In this analysis, we do not consider how the levels are formed.

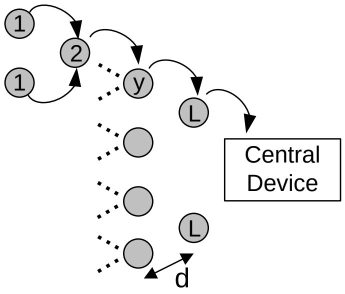


Figure 3.6: Example of a balanced tree topology with ζ = 2 for the evaluation of multi-hop or single-hop architecture.

Similar to the line topology analysis, we will compare the ratio of energy consumption for multi-hop, i.e. when the tree structure is used and for single-hop communication. Assume that there are L levels in the network. The energy usage for a packet of k bits for a node at level y when using single-hop can be written as

$$E_{SH}(y,d) = E_{TXelec} \cdot k + E_{amp}(\eta)\left((L-y+1)\cdot d\right)^{\eta} \cdot k \qquad (3.6)$$

Whereas the energy usage for a node at level y in a multi-hop network is given as:

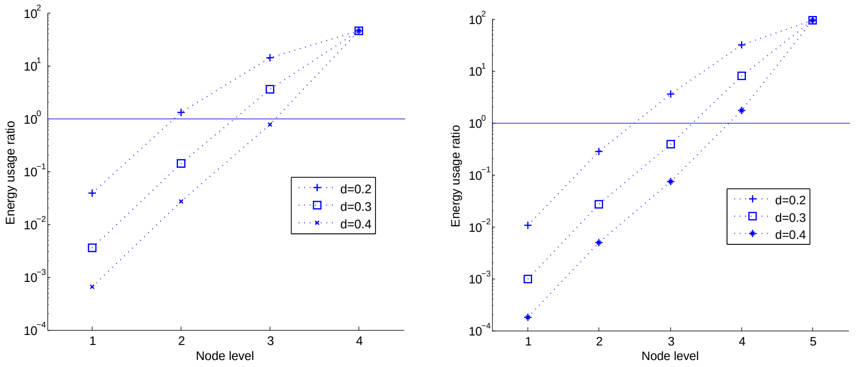
$$E_{MH}(y,d) = \zeta \cdot \sum_{i=0}^{y-2} \zeta^{i} \cdot E_{RXelec} \cdot k + \sum_{i=0}^{y-1} \zeta^{i} \cdot k \cdot (E_{TXelec} + E_{amp}(\eta) \cdot d^{\eta}) \quad (3.7)$$

Each node now has ζ children due to the ζ-balanced tree, what explains the factor ζ at the receiver part. If ζ = 1, the formula reduces to (3.4). Thus, the line topology can be seen as a special case of the tree topology.

As in the line topology, the higher path loss exponent is used when the nodes are not neighboring. The performance ratios are plotted in Figure 3.7 for ζ = 2. The situation is quite similar to Figure 3.4. The tree topology and the resulting higher forwarding overhead makes the nodes near the sink perform even worse, further away from the sink the single-hop situation remains the same.

When aggregation is used, the energy used for multi-hop by a node on level y is given by

$$E^{A}_{MH}(y,d) = k \cdot \zeta \cdot \sum_{i=0}^{y-2} (\zeta \cdot \phi)^{i} \cdot E_{RXelec} + k \cdot \sum_{i=0}^{y-1} (\zeta \cdot \phi)^{i} \cdot (E_{TXelec} + E_{amp}(\eta) \cdot d^{\eta}) \qquad (3.8)$$



(a) Tree with 4 levels

(b) Tree with 5 levels

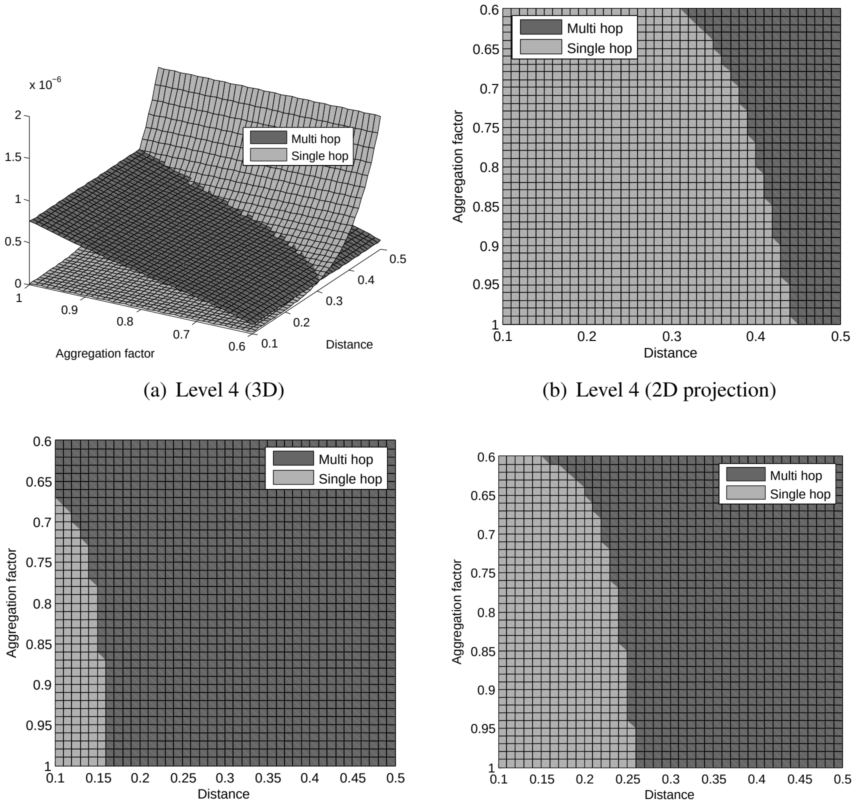
Figure 3.7: Comparison of the energy usage in a single-hop and multi-hop in a scenario with a balanced tree consisting of 4 and 5 levels and $\zeta = 2$ (Fig. 3.6). The ratio of the energy usage of multi-hop and single-hop communication is given (E_{MH}/E_{SH}). When the ratio is higher than 1, single-hop is the most energy efficient.

The proof of the summation is similar to the one of the line topology. If $\phi = 1$, the formula reduces to (3.7).

Figure 3.8 shows the most optimal architecture for a tree topology of 5 levels and $\zeta = 2$ for the nodes on levels 2, 3 and 4. Figure 3.8(a) gives a 3D-plot of the energy consumption for the multi-hop and single-hop scenario for varying ϕ and d for level 4. A 2D-projection is given in Figure 3.8(b). It can be seen that the influence of the aggregation factor is larger in a tree network as the nodes have more children and consequently more data can be aggregated. Aggregation is thus more useful in tree networks.

In Figure 3.9 we have plotted the most optimal architecture for a balanced tree with $\zeta = 3$ and $\zeta = 4$ for the nodes on level 3. We can see that, when no aggregation is used, the more children a node has, the further the nodes should be placed apart before a multi-hop topology becomes more energy efficient. This is expected as the nodes now have to relay more data because the increased number of children. This induces an overhead with regard to the energy consumption. However, as can be noticed in the figures, when aggregation is used, the distance between the nodes becomes less important. It can therefore be concluded that aggregation is certainly useful in tree networks.

We have investigated the most energy efficient architecture for two simple topologies: all nodes on one line and a binary tree. We have found that whether to use a multi-hop or single-hop approach depends primarily on the distance between the nodes and between the nodes and the sink. If we look at both the line



(a) Level 4 (3D)

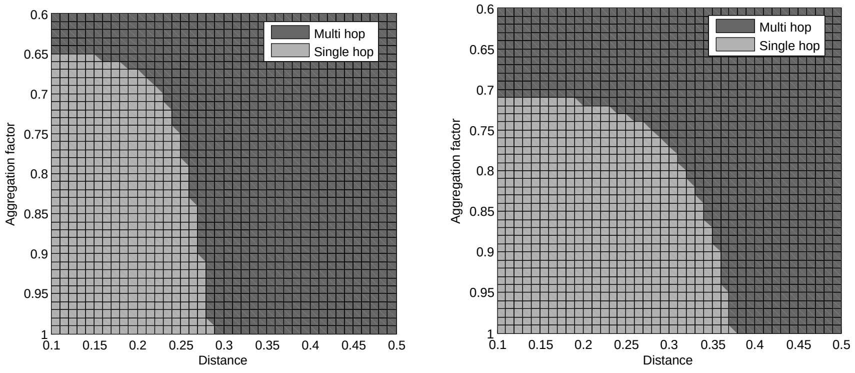
(b) Level 4 (2D projection)

(c) Level 2 (2D projection)

(d) Level 3 (2D projection)

Figure 3.8: Influence of data aggregation and distance on the energy consumption for a tree topology of 5 levels and $\zeta = 2$ (Fig. 3.6) for the nodes on level 2, 3 and 4. The lighter area means that a single-hop topology is more energy efficient. The distance is expressed in meters [m].

and the tree topology, we see that in single-hop there is clearly room for energy saving at the nodes further away from the sink. These nodes consume the most energy and consequently will die first. However, we also see that in the multi-hop scenario, more energy is consumed by the nodes closest to the sink as they have to forward the data received from nodes farther away. When a tree topology is used, aggregation also plays an important role. As this case is more general, we will try to improve performance for this scenario. Any improvements will then be trivially the same for the single-hop scenario. In the current evaluation, we have only considered communication between direct neighbors or with the personal device. The most optimal architecture will use a mixture of single and multi-hop: when



(a) $\zeta = 3$

(b) $\zeta = 4$

Figure 3.9: Influence of data aggregation and distance on the energy consumption for rising ζ for the nodes on level 3. The lighter area means that a single-hop topology is more energy efficient. The distance is expressed in meters [m].

the nodes are close, communication will be single-hop. This will be studied in the following section.

3.3 Cooperation

In this section, we will expand the research to more general network architectures and more possible solutions. Instead of only considering multi-hop and single-hop, we will now look into a mixture of both architectures where a few nodes send their data using multi-hop and other nodes, directly to the personal device. This approach is called *cooperation*. The concept of cooperation is made possible by the broadcast nature of the radio channel. The basic idea is that one-hop neighbors of a transmitting node can overhear this transmission and may offer their cooperation in making the information reach the destination. This leads to a considerable network performance increase as shown in [14] and [15]. The overall goal is to maximize the lifetime of the network, where the lifetime is defined as the time until the first node runs out of energy.

The problem of minimizing energy cost and maximizing lifetime has been considered before for sensor networks. In [6, 16] the upper bound of the lifetime of the sensor network is determined and in [17] the problem is extended to a scenario where a few nodes can fall out as long as the connectivity is respected. In [18] an ILP formulation is used to determine the most energy efficient routing scheme. It is concluded that the ratio of circuit and transmission energy needs to be considered.

We will first illustrate this type of cooperation in Body Area Networks by an

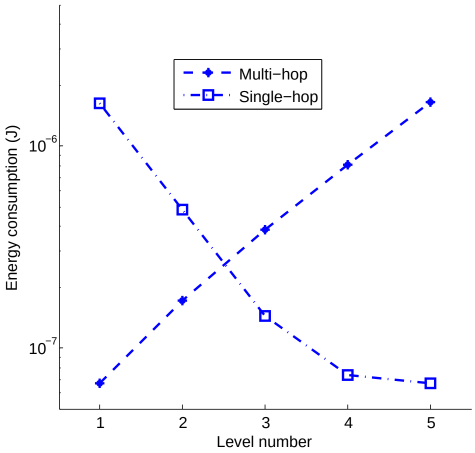


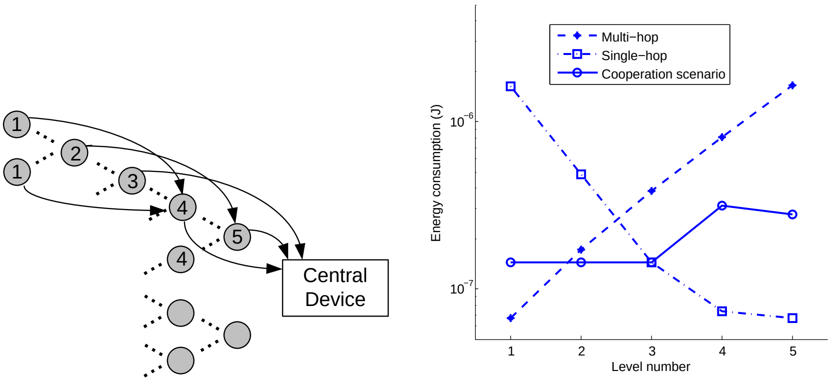
Figure 3.10: Energy comparison of single-hop and multi-hop in a binary tree network with 5 levels for nodes on level 1 to 5.

example. The problem is subsequently represented as an ILP formulation. Three objectives are considered: minimize the overall network energy usage, minimize the maximum energy usage per node, i.e. maximize the lifetime of the network and minimize the number of hops with the maximum energy usage as constraint.

3.3.1 Example

Looking at Figure 3.7(a) and 3.7(b), it is clear that for levels 4 and 5 the single-hop architecture consumes a lot less energy than when multi-hop is used. Figure 3.10 explicitly plots the energy consumption in single-hop and multi-hop. In the single-hop architecture, nodes 1 and 2 consume considerable more energy when they send their data directly to the central device. The proposed solution is to use the residual energy of the nodes at level 4 and 5 for relaying data from other nodes. Stated otherwise, to let those nodes cooperate in the network. Indeed, by breaking into this energy supply, the lifetime of the network will be higher as the advantages of a multi-hop network for level 1 and level 2 nodes is combined with the available energy at level 4 and 5. The data of the nodes on level 1 and 2 can be forwarded to level 4 and level 5 respectively. Thus, in stead of only forwarding data to the most nearby node, the data is transported more intelligently and the burden of forwarding the data is more equally spread among the nodes. The architecture no longer only considers the naive approach of only sending data to a node's direct neighbors.

An example of the resulting network topology with communication links is shown on Figure 3.11(a). Figure 3.11(b) shows the energy consumption when using this approach, compared with multi-hop and single-hop scenarios. The energy consumption of a node at level 4 or 5 sending its own data and relaying the data of the other nodes, remains below the energy consumption of the multi-hop scenario. Thus, the lifetime of the network is improved. The almost horizontal energy usage line demonstrates a good trade off between the peeks when using a smart combination of simple single-hop and multi-hop network setups. The cooperating network lifetime is a lot higher compared to the single-hop or multi-hop approaches of figure 3.10 as the maximum energy is lower. In the full binary tree structure of this example, we relay 8 nodes on level 4 from level 1 and 8 nodes on level 5 from level 2.



(a) Example of a cooperating tree topology. The arrows indicate the communication links.

(b) Energy usage when cooperating is used compared with single-hop and multi-hop energy consumption

Figure 3.11: Tree topology with cooperation and resulting energy usage.

It is interesting to calculate how much data can be forwarded by a node i, i.e. the number of nodes that can use node i as a relay station towards the sink. Based on (3.6) and (3.7), the following formula can be used for a node on level k when the energy consumption is limited by an energy E_{lim}:

$$\#\text{nodes supported} = \left\lfloor \frac{E_{lim} - E_{SH}(k,d)}{E((L-k+1)\cdot d)} \right\rfloor \tag{3.9}$$

where $E(d)$ is defined as the energy needed for receiving data of one node and relaying it over a distance d. The denominator thus expresses the energy needed for relaying data from one other node. The numerator calculates the residual energy that is available in the node on level k for relaying if the available energy is limited to E_{lim}. Consider as example the network of Figure 3.11. We have taken the energy consumption for multi-hop at level 4 as restriction (i.e. $E_{lim} = E_{MH}(4,d)$). This means that a node on level 4 can relay data from up to 12 nodes and a node on level 5 data from up to 13 nodes.

The analysis for a line topology is similar and can be seen as a special case of the tree topology where the number of children is restricted to 1.

3.3.2 Network Model

A wireless Body Area Network can be modeled as a directed graph G where V is the set of nodes including the sink or personal device. The sink is represented as S. A unidirectional link exists between two nodes m and n if n is within the send range of m and m is within the receive range of n. The number of nodes is N. V_s is the set of nodes without the sink S ($V_S = V \backslash \{S\}$). The distance between nodes i and j is $d_{i,j}$.

The variables of the optimization problem:

$$y_{i,j}^{k,l} = \begin{cases} 1, \text{node } i \text{ relays a packet to node } j \\ \text{ with origin } k \text{ and destination } l \\ 0, \text{otherwise} \end{cases} \tag{3.10}$$

As all the data is sent to the sink, the variable is simplified to $y_{i,j}^{k}$. One can argue to even simplify the variable further by not considering the origin of the packet. We choose not to do so, as we can use this information to rapidly determine how the information of node k is forwarded to the sink, e.g. to count the number of traversed hops or to calculate the path probability, as used in Section 3.6.

The total transmission energy of node $i \in V_s$ can then be calculated using the radio model of Section 2.3.3.2:

$$E_{TX}^{i} = \sum_{j \in V} \sum_{k \in V_S} y_{i,j}^{k} \cdot [E_{TXelec} + E_{amp}(\eta) \cdot (d_{i,j})^{\eta}] \tag{3.11}$$

The receive energy of node i only accounts for the energy consumed by the circuitry and is given as

$$E_{RX}^{i} = \sum_{j \in V} \sum_{k \in V_S} y_{j,i}^{k} \cdot E_{RXelec} \tag{3.12}$$

The energy required for sensing is represented as E_{sense} and is the same for all sensing nodes. A typical value is 50nJ/bit [19]. Aggregation of data is not taken into account.

3.3.3 Constraints

The following constraints are used:

- Every node has a packet generation rate of 1:

$$\sum_{j \in V} y^{i}_{i,j} \;=\; 1;\; \forall\, i \in V_s \tag{3.13}$$

- Every node forwards all the data it receives:

$$\sum_{j \in V} y^{k}_{i,j} \;-\; \sum_{j \in V} y^{k}_{j,i} \;=\; 0;\; \forall\, i, k \,\in V_s;\; i \neq k \tag{3.14}$$

  The first part accounts for the forwarding of the data originated at node k by node i to all possible nodes. The own generated data is not taken into account ($i \neq k$), as this is covered by (3.13). The second part reflects the reception of data originating at node k by node i, received from all possible nodes.

- The sink receives all the data:

$$N \,-\, 1 \,=\, \sum_{k,j \in V_s} y^{k}_{j,S} \tag{3.15}$$

3.3.4 Objectives

We consider different objectives. In the first, the overall energy usage over the network is minimized (`total energy`):

$$\text{Obj: minimize} \sum_{i \in V_s} \left[E^{i}_{RX} + E^{i}_{TX} + E_{sense}\right] \tag{3.16}$$

This objective does not take into account the individual energy usage of the nodes. It might therefore well happen that one node has a very high energy consumption compared to the others. This will lead to a low network lifetime. For the second objective, we will consider the individual energy consumption and as a consequence maximize the lifetime of the network. This can be achieved by finding the minimum of the maximum node energy usage. However, this approach only considers the energy consumption of the most consuming node and does not influence the energy usage of the other nodes. Some nodes may thus route their packets inefficiently. Therefore, a low number of hops is still desired and an ILP consisting of 2 phases is used. In the first phase, the maximum node energy usage is minimized (`max energy`):

$$\text{Obj: minimize} \max_{i \in V_s} \left[E^{i}_{RX} + E^{i}_{TX} + E_{sense}\right] \tag{3.17}$$

In a second phase, this objective value (E_{max}) is used as a maximum value for the energy consumption and the new objective is to minimize the number of hops (`connections`):

$$E_{RX}^{i} + E_{TX}^{i} + E_{sense} \; < \; E_{max}; \; \forall \, i \in V_s \tag{3.18}$$

$$\text{Obj: minimize} \sum_{i,j \in V} \sum_{k \in V_S} y_{i,j}^{k} \tag{3.19}$$

3.3.5 Analysis

3.3.5.1 Setup

All integer linear programs needed to evaluate the cooperation problem, have been solved using ILOG CPLEX 10.0 [20], running on BEgrid, the computing/data grid infrastructure that results from the Belnet Grid Initiative [21].

The number of nodes is varied from 5 to 15. We have evaluated the average maximum network energy consumption in Joule (J) (Figure 3.12) which is a criterion for the network lifetime and the average number of hops per connection (Figure 3.13). The networks are randomly generated in an area of 2m by 2m and the personal device is placed in the middle. Each data point represents an average of 100 runs.

3.3.5.2 Discussion

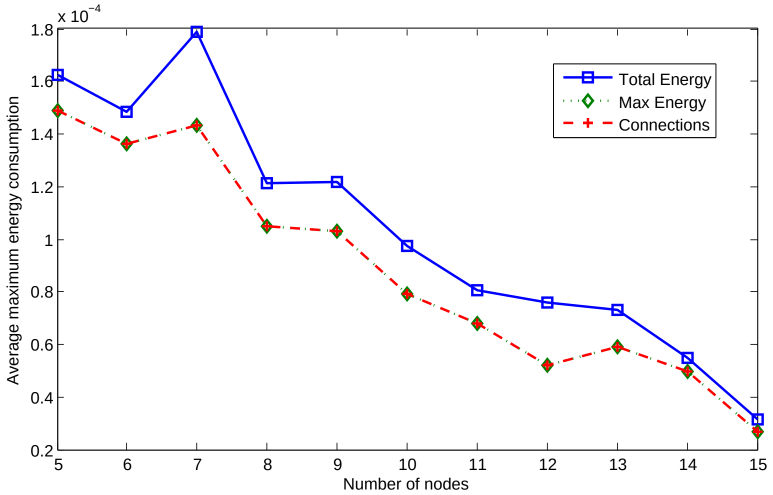


Figure 3.12: The maximum node energy consumption for varying number of nodes for the different objectives: `total energy` (3.16), `max energy` (3.17) and `connections` (3.19). The graphs for `max energy` and `connections` coincide due to restriction (3.18). For a higher network lifetime, it is better to consider the maximum node energy.

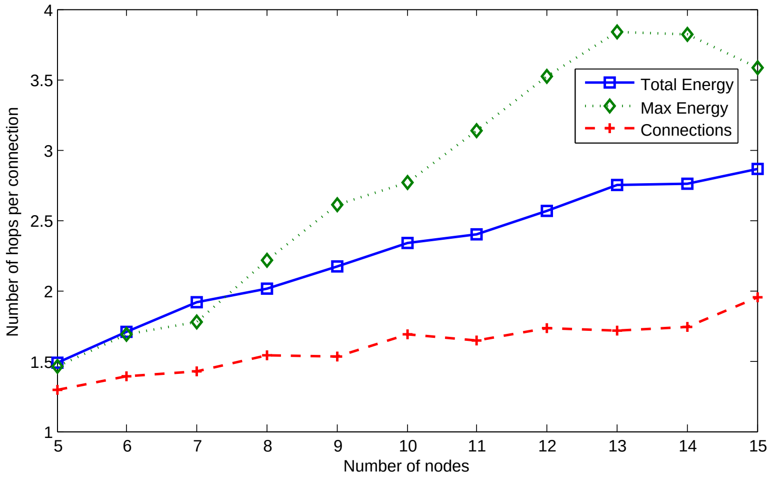


Figure 3.13: The average number of hops per node for a varying number of nodes for the different objectives: `total energy` (3.16), `max energy` (3.17) and `connections` (3.19). It can be seen that the average number of hops is less than 2 when the number of connections is minimized.

The lowest lifetime is achieved when the overall energy usage is minimized (`total energy`). This is to be expected as the energy usage of the individual nodes is not considered and one node can have a very high load. When the second objective is used (3.17), the lifetime of the network is improved, but a larger number of hops is used. This was expected as the ILP-solution only looks at the node with the maximum lifetime. Thus, by using the extra condition (3.18) and objective function (3.19), the number of hops is minimized whilst preserving the lifetime, as can be seen in Figure 3.12 (`connections`). So, the use of these last objectives leads to an optimal network with a minimum number of hops and a maximum lifetime of the network. The lifetime of the network is increased by 24% on average.

The maximum network energy consumption lowers when the network has more nodes. The more nodes available, the more the burden of forwarding data will be distributed, which leads to a decrease of the energy consumption. Further, even when the number of hops per connection is minimized, this is still between 1.5 and 2. This means that forwarding data from other nodes is beneficial for the lifetime of the network. This is in line with (3.2) where d_{char} is 43 cm and the maximum distance between a node and the personal device is 1 m, giving a K_{opt} of 3. As the nodes are uniformly distributed in the area, this is averaged out to 1.5. Moreover, the use of objective (3.19) has the side effect that a node will send his data directly to the personal device as long as its energy consumption is lower than the maximum energy consumption. This also lowers the number of hops per connection.

3.4 Relaying

The cooperation from the previous section, proves to be very useful. But the energy consumption can be lowered even more by introducing dedicated relay devices in the network. These are special nodes which only handle traffic relaying and do not do any sensing themselves, thus more energy is available for communication purposes. The main idea is that proper placement of relay nodes can bridge the performance gap for the nodes far away from the sink in the case of single-hop traffic and offload the nodes closer to the sink in the case of multi-hop traffic. Further, the relay devices are dedicated devices with no sensing functionality that merely forward data received from other nodes. This means that more of the available energy can be used for forwarding and receiving data. The sensor device handles the received data the same way as a normal node does: the received signal is decoded, when the signal is not properly received, a retransmission is done. Doing so, propagation of errors in the network is avoided.

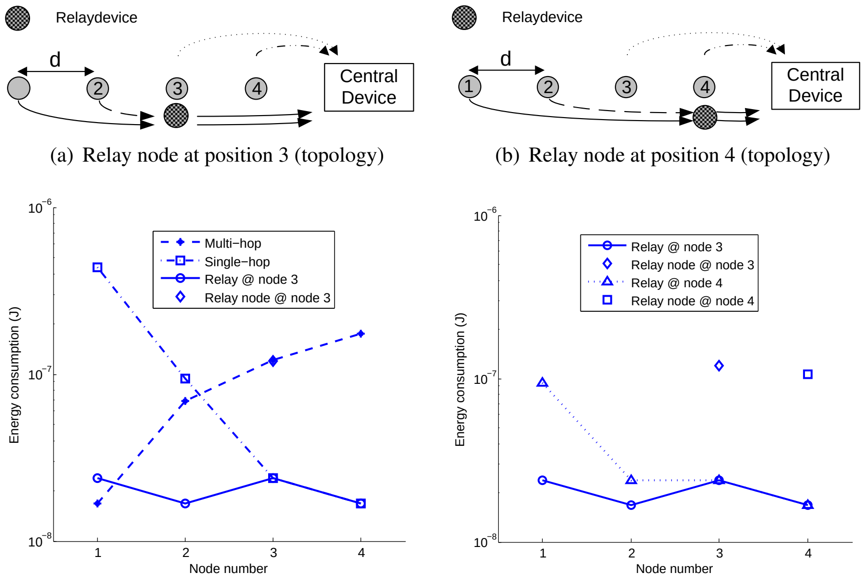
For a node relaying traffic from z nodes, energy usage for packets of k bits is similar to a regular node in a multi-hop network (3.4), minus the cost of transmitting own packets:

$$E_R(z,d) \;=\; z \cdot k \cdot E_{RXelec} + z \cdot k \cdot \phi \cdot (E_{TXelec} + E_{amp}(\eta) \cdot d^{\eta}) \tag{3.20}$$

In this formula, d represents the distance to the next relay hop. Further, this equation is independent of the architecture used, i.e. line or tree topology. However, the number of nodes that are relayed, z, depends on the position of the relay node and the architecture used. As in the previous analysis, aggregation with a factor ϕ can be used.

3.4.1 Line Topology

We will evaluate the use of relay devices for the line topology of Figure 3.3(b). One extra relay device is used and is placed along the position of node 3 or node 4. Nodes 1 and 2 forward their data to the relay device (Figs. 3.14(a) and 3.14(b)). The resulting energy usage can be seen in Figure 3.14. Figure 3.14(c) shows for each node the usage when using the single-hop, the multi-hop and the relay scenario. It can be seen that the nodes consume the least in the relay node scenario. The energy consumption of the relay device has also been plotted. It is clear that the relay device will be the limiting factor of the lifetime of the network. By moving the position of the node along node 4, the energy consumption of the relay device decreases, see Figure 3.14(d).



(a) Relay node at position 3 (topology)

(b) Relay node at position 4 (topology)

(c) Relay node at position 3 vs single-hop and multi-hop

(d) Relay node at position 3 vs position 4

Figure 3.14: Energy usage when relaying in a line topology. No aggregation is used (ϕ = 1). The energy consumption of the relay devices is also plotted (relay node position)

When aggregation is used by the relay device, the energy consumption will decrease further as less bits need to be sent.

3.4.2 Tree Topology

As an example, consider a binary tree network of 5 levels where the relay devices are placed at level 4, see Figure 3.15. Each node at level 4 has a relay device next to it. The relay devices only forward data from the nodes at levels 1 and 2 and relay it directly to the sink. Figure 3.16(a) shows the result when the distance between the nodes is 20 cm and figure 3.16(b) gives the result for a network with 30 cm between the nodes. The graphs for single-hop and multi-hop communication are plotted. Further, the energy consumption of the relayed network is shown. It can be seen that the lifetime of the nodes at level 1 and 2 improves a lot with respect to the single-hop scenario. In both cases the energy usage at level 1 decreases by a factor 10.

The points at level 4 represent the energy usage of a relay device when it forwards data from 12 or 6 nodes. When 12 nodes are forwarded, the energy con-

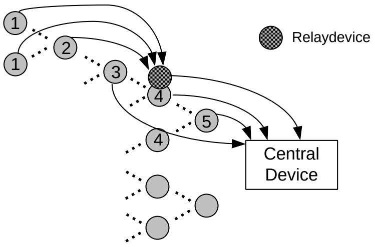
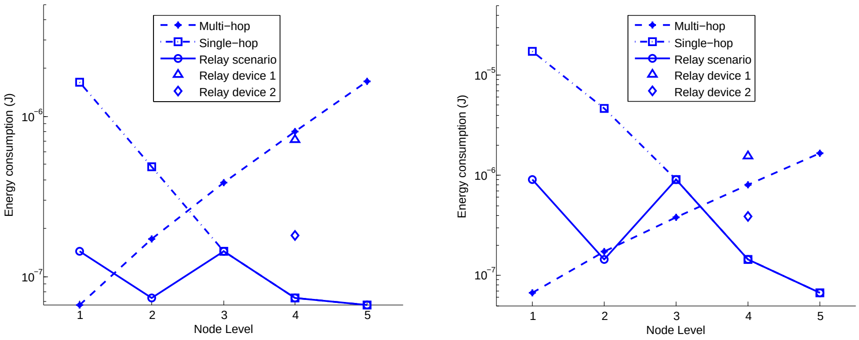


Figure 3.15: Example of a balanced tree topology with ζ = 2 when a relay device is added on level 4. The communication links are shown on the figure. No aggregation is used (ϕ = 1)



(a) Inter node distance = 20 cm

(b) Inter node distance = 30 cm

Figure 3.16: Energy usage when relaying in a tree topology for varying inter-node distances and ζ = 2 (Figure 3.15). No aggregation is used (ϕ = 1). Relay device 1 forwards data from 12 nodes, relay device 2 forwards data from 6 nodes. The energy consumption of the relay devices is also plotted.

sumption of the relay device is about 6 times higher than the one of the sensor nodes in the relay scenario. When an extra relay device is added, the energy consumption of the relay devices is only slightly higher than the maximum energy consumption of the sensor nodes in the relay scenario. When the distance between the nodes is 20 cm (Figure 3.16(a)), we see that the energy needed for relaying data of three nodes is lower than the energy usage of the nodes at level 3. If the distance is larger, i.e. 30 cm, it is even possible to relay 7 nodes while staying under the energy consumption at level 3 (Figure 3.16(b)). The exact number of nodes whose data can be forwarded can be calculated by a formula similar to (3.9). But as the relay device has no data to send, the formula reduces to

$$\#\text{nodes supported} = \left\lfloor \frac{E_{lim}}{E_R((L-k+1)\cdot d)} \right\rfloor . \tag{3.21}$$

with L the number of levels of the tree and k the level where the relay device is added and E_R is (3.20). Depending on the energy required for sensing, supporting a larger number of nodes should be possible.

It should be noted that the number of nodes in this example network is very high, the number of relay nodes will not have to be that high in realistic WBANs.

Figure 3.17 shows the influence of data aggregation when relaying is used and when the number of relayed nodes varies. The aggregation of data lowers the energy consumption.

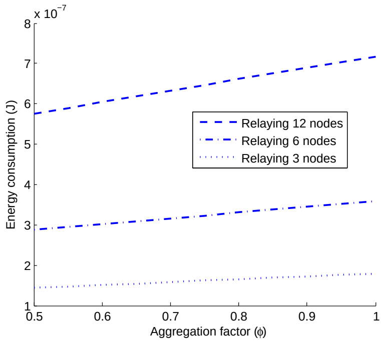


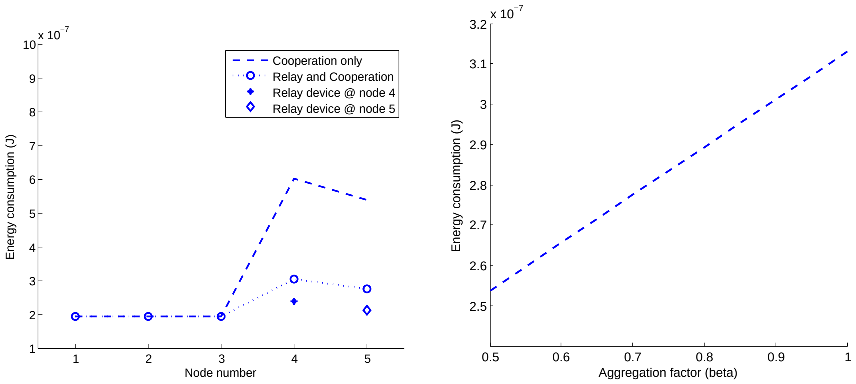
Figure 3.17: Energy usage for a relay device placed at level 4 (Figure 3.15) when relaying in a tree topology for varying number of relayed nodes and aggregation.

When considering networks with more hops, the introduction of relay devices clearly shows a better performance. This is caused by the high path loss around the body. The position of the relay nodes is highly situation dependent. Yet, the following rule of thumb can be used: the placement should not be too far away from the sink as the path loss effects will impact efficiency dramatically. A position closer to the sink is a better option, however the number of hops between the nodes and the relay device should not become too large.

3.5 Relaying and Cooperation

The previous section shows that using relay nodes considerably improves the lifetime of the network. However, it is not always feasible to use relay nodes or to add an unlimited number of relay nodes. Specifically in the case of Body Area Networks, putting even more sensor nodes on users does not really improve comfort. Therefore, we will use a combination of adding relay devices and cooperation of other nodes. A similar approach was presented recently in [22] for wireless networks where a combination of cooperation and relaying is used. This strengthened our belief that this approach is the proper way to go.

As an example, consider the cooperating tree of Figure 3.11. The most energy is consumed by nodes at levels 4 and 5. We now can lower this maximum by adding relay devices at level 4 and 5. The resulting energy consumption is shown on Figure 3.18(a). The energy consumption of the relay devices is also shown in the figure. For the nodes, the energy consumption has decreased with 43% and the network lifetime has risen accordingly. When the data is aggregated before sending it, it is clear that the energy consumption will decrease even further, as can be seen in Figure 3.18(b) that plots the maximum energy consumption of the nodes and relay devices. When ϕ = 0.5, an extra lifetime increase of 18.3 % is achieved.



(a) Overview per node. ϕ = 1 (no aggregation used)

(b) Influence of aggregation. The maximum energy consumption of the nodes and relay devices is plotted.

Figure 3.18: Energy consumption of a tree topology with ζ = 2 using relay and cooperation, see Figure 3.11. Relay devices are added next to the nodes at level 4 and 5.

3.5.1 Adding Relaying to the ILP

3.5.1.1 Network Model

As in section 3.3, the problem of relaying and cooperation will be formulated as an ILP. For simplicity, it is assumed that the relay devices are placed in addition to and next to the regular nodes (i.e. the nodes sending data). Let V_s be the set of regular data nodes without the sink and V_r the set of relay devices. In all equations, replace V_s with $V_r \cup V_s$, except in (3.13) as the relay devices do not generate data.

Add the following constraint: a relay device does not generate data:

$$\sum_{j \in V} y^i_{i,j} = 0; \ \forall i \in V_r \tag{3.22}$$

The relay devices do not consume energy for sensing, so they can spend a larger amount of battery energy for forwarding data. Thus, for all relay devices the sensing energy E_{sense} is set to zero.

A next step is to restrict the number of relay devices, as this could increase the ease of use. This is done by adding the restriction that a relay device only can be added when it has a minimum of 4 links. This means that the data of at least 2 nodes is received and forwarded by the relay devices.

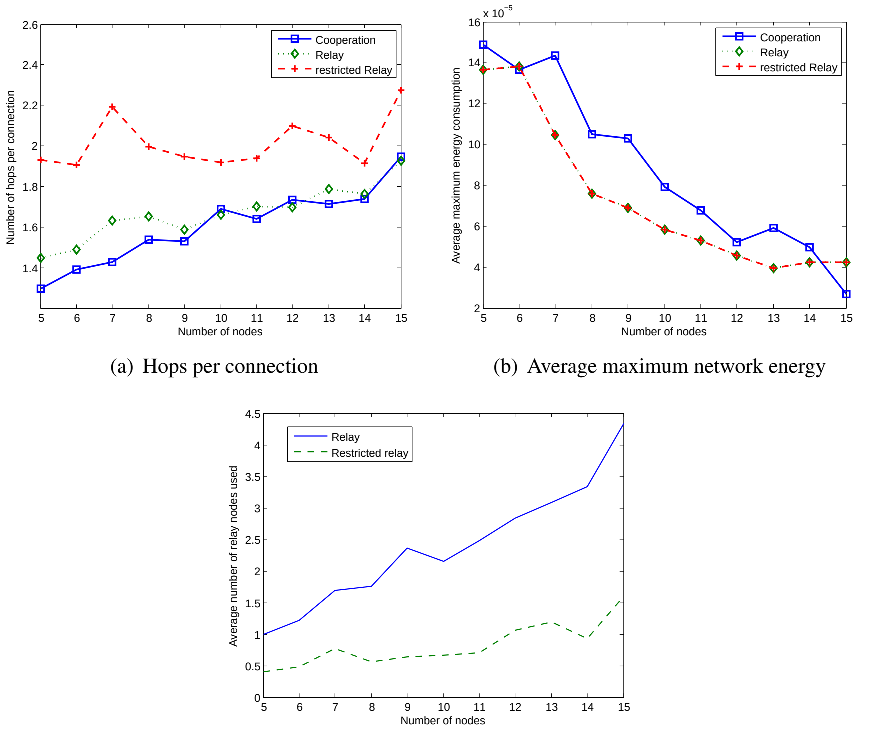
$$\sum_{j,k \in V} y_{i,j}^{k} > 4 \vee \sum_{j,k \in V} y_{i,j}^{k} = 0; \; \forall i \in V_r \tag{3.23}$$

The objectives remain the same: first the maximum network energy consumption is minimized and in a second stage the number of hops per connection is minimized.

3.5.1.2 Evaluation

The number of normal nodes is once again varied between 5 and 15 and the same number of (possible) relay devices as the number of nodes in the network is used. The results with and without relay devices are compared for objective (3.19) only, i.e. the number of hops is minimized (`connections`). `Cooperation` refers to the scenario without relay devices, `restricted relay` uses constraint (3.23). It can be seen that the average number of hops per connection (i.e. per node generating data) is almost the same, see Figure 3.19(a), only a little bit higher. This is normal as in this scenario some nodes will send their data to a relay node instead of sending it directly to the central device. In small networks, less hops are needed when no relay devices are used. This means that not all the nodes use the relay devices. When the number of nodes is restricted, the number of hops needed has increased as less relay devices will be used. The most interesting results can be seen on Figure 3.19(b), as it shows that the maximum network energy is lower when relay devices are used. Lifetime increases up to 30% are reached, depending on the number of nodes used. When the number of relay devices is restricted, the lifetime is the same. The average number of relay devices used is lower than 1 when the relay devices are restricted. This means that not in every topology a relay device is needed. The lifetime of the network is the same whether or not the number of relay devices is restricted. This is because of contstraint (3.18) that limits the maximum network energy consumption. As WBANs most likely will consist out of more than 10 nodes, this analysis shows that using relay devices is useful for increasing the lifetime of the network. Figure 3.19(c) shows the number of relay devices that are used.

The analysis and discussion of this section has shown that combining relaying, i.e. adding extra relay devices, and cooperation of the other nodes is beneficial for the lifetime of the network. The ILP clearly indicates this benefit for larger



(a) Hops per connection

(b) Average maximum network energy

(c) Average number of relay devices used

Figure 3.19: ILP with relaying. The restricted relaying is where the number of relaying nodes is restricted (constraint (3.23)) . For all scenarios, the network lifetime was maximized and the number of hops minimized, see objective (3.19).

networks especially. Further, when the number of relay devices is limited, the maximum network energy consumption is even lowered more. The ILP formulation has not considered aggregation. It is expected that using aggregation will be even more energy efficient.

3.6 Reliability

While the previous sections looked into the energy consumption of a Wireless Body Area Network, this section will focus on the reliability of the communication. Reliability is very important in a Body Area Network, as we are handling delicate medical data and loss of data needs to be minimized, see Section 2.1.4. Several metrics exist to express the reliability: the packet delivery ratio, the loss probability, probability of collision and so on. These metrics and their conse-

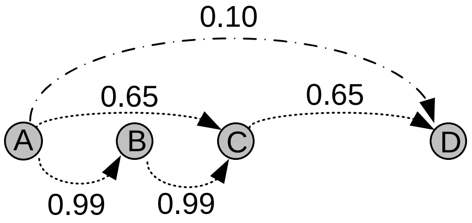


Figure 3.20: Example of a connection probability in a single-hop and a multi-hop scenario. The distance between node A and B is 10 cm and between node C and D 20 cm. Multi-hop communication between node A and D is more reliable than single-hop communication.

quences on the network topology are discussed in the following. We do not consider the delay in this study, this will be subject of future research.

3.6.1 Connection Probability in Multi-hop Networking

In Section 2.3.2 the probability of a link in a WBAN was derived. Using this probability, the reliability of the communication can be derived and compared for a single-hop or a multi-hop topology.

In order to develop an intuition why there might be room for improvement in multi-hop routing, it is helpful to consider Figure 3.20. Different nodes are placed on one line and different routes are shown for communication between nodes A and D. The numbers above the communication links show the link probability between the two nodes using (2.12) and the variables of Table 2.3. At one extreme, node A could send directly to D in one hop and at the other extreme, A could use the 3-hop route through B and C. In the example, it is clear that the 3-hop communication has a communication probability of 63.7% whereas the single-hop communication is only 10%. On the other hand, in multi-hop communication nodes C and D will hear many of the packets send from A to B and it is wastefull that node B forwards these packets. This example shows the trade-off between the reliability and the energy efficiency.

Using the formula for link probability $p(d)$ (2.12), we can derive the following condition to determine whether or not multi-hop communication should be used in terms of reliability. Assume that two nodes are sending to each other and are placed a distance d apart. If n nodes are placed equidistantly between these nodes, then the condition to use multi-hop communication becomes:

$$p\left(\frac{d}{n+1}\right)^{(n+1)} > p(d). \tag{3.24}$$

When applying this inequality for the link probabilities plotted in Figure 2.5, it turns out that the multi-hop path has the highest reliability, independent of the distance d or the number of intermediate nodes n. This can be proved using the following proposition:

Proposition 1. *If $n > 1$ and $0 < x < 1$, then the following inequality holds:*

$$x^n \; > \; 1 \, + \, n \cdot (x - 1)$$

Proof:

We use the following identity:

$$\frac{x^n - 1}{x - 1} \; = \; \sum_{i=0}^{n-1} x^i$$

For $x < 1$, the terms in the right hand side can not be higher than 1. This means that the sum on the right hand side can not be higher than n as we have n terms. We therefore get

$$\frac{x^n - 1}{x - 1} \; < \; n$$

Taken into account that for $x < 1$, the denominator is negative, we get the proposition. □

Thus, using this proposition and the fact that the probability is between 0 and 1, we can say

$$p\left(\frac{d}{n+1}\right)^{(n+1)} \; > \; 1 \, + \, n \cdot \left(p\left(\frac{d}{n+1}\right) - 1\right) \tag{3.25}$$

Further, we know that $p\left(\frac{d}{n+1}\right) > p(d)$ as the link probability is a monotonically decreasing function (see Figure 2.5). Thus we can write:

$$p\left(\frac{d}{n+1}\right) - 1 \; > \; p(d) - 1$$

$$\overset{n \geq 1}{\Rightarrow} \; n \cdot \left(p\left(\frac{d}{n+1}\right) - 1\right) \; > \; p(d) - 1$$

Combining this with (3.25), we have proved the correctness of (3.24).

One has to keep in mind that this proof only holds as long as the intermediate hops are placed equidistantly on the path between the sender and destination. Further, as the probability is a statistical value this will not always be the case. At a given moment in time, the reliability over the multi-hop path can be lower as for example the path between node C and D may experience high packet loss temporarily. Of course, when nodes A and D are sufficiently close to each other, the reliability of direct communication will be high enough to use it and the gain of using multi-hop communication will be negligible.

The connection probability $CP_{i,j}$ between nodes i and j depends on the route between the nodes. For a path r, it can be written as

$$CP_{i,j}^{r} = \prod_{l \in r} p(d_l), \tag{3.26}$$

where $p(d_l)$ is the link connectivity of link l and d_l the length of link l. The most optimal path across all routes r from i to j is then given by

$$CP_{i,j} = \max_{r} CP_{i,j}^{r}. \tag{3.27}$$

The connection probability can be used as a measure for the packet delivery ratio when a perfect duty cycle is assumed, i.e. when no collisions occur. It approximates the end-to-end reliability of a routing path in the absence of retransmission [13].

Overall, it shows that a multi-hop architecture has the highest reliability. This result is in line with the conclusions of [23] and the study in [10] where it is experimentally shown that the packet delivery ratio can be improved by using a multi-hop architecture.

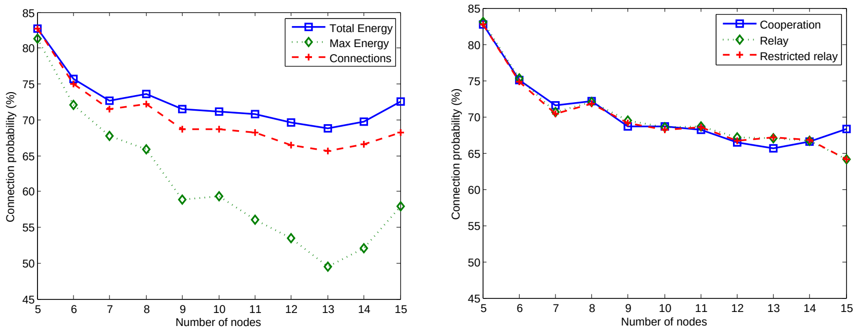
3.6.2 Reliability in ILP

In this scenario, we will investigate the influence of adding relay devices or cooperation on the reliability of the connections for the ILP formulation of Section 3.3. The following formula expresses the connection probability of the data sent by node i to the sink.

$$CP_{i,S} = \prod_{(j,k) \in A_i} p(d_{j,k}) \tag{3.28}$$

where $A_i = \{(j,k) | j,k \in V \wedge y_{j,k}^{i} = 1\}$ or the set of edges on the path from node i to the personal device S.

Figure 3.21 shows the connection probability for the scenarios with and without the relay devices for the following objectives: sum of overall energy is minimized (`Total Energy`) (3.16), maximum network consumption is minimized (`Max Energy`) (3.17) and the number of hops is minimized (`Connections`) (3.19). It can be seen that the highest reliability is obtained when the overall network energy is minimized. The worst reliability is obtained when only the lifetime of the network is minimized. Indeed, this objective leads to longer paths with a higher number of hops per connection, see Figure 3.13. These paths are not optimized in any way and are therefore more likely to contain links with high loss



(a) Without relay devices

(b) With relay devices

Figure 3.21: Reliability in the ILP for the scenarios with and without relaying.

rate. When the number of hops is minimized, the reliability improves. The reliability is lower than in the first case as more nodes now send their data directly to the personal device as long as their energy consumption is below the maximum network energy. The addition of relay devices does not influence the reliability, see Figure 3.21(b). The number of hops per path is a little bit higher when relay devices are used. The reliability is about the same as nodes now send their data to a relay device instead of to the sensor node located next to the relay device. This communication has the same link probability and therefore the connection probability, i.e. the reliability, remains the same.

Now we will try to improve the reliability by adding a new objective to the ILP-formulation which can be used instead of minimizing the number of hops (3.19). We do not longer minimize the number of hops, but maximize the reliability. This can be done by maximizing the minimum of (3.28) over all connections, but a product can not be used in a linear integer formulation. Therefore, we will transform (3.28) using the natural logarithm. As this is a continuous rising function, the minimum or maximum will not change. (3.32) can then be written for a given node i as

$$\ln\left[\prod_{(j,k)\in A_i} p(d_{j,k})\right] = \sum_{(j,k)\in A_i} \ln(p(d_{j,k})) \tag{3.29}$$

where A_i is the set of edges on the path from node i to the personal device S. Using the definition of $y^i_{j,k}$ in (3.10), the summation over the edges in A_i can be written as:

$$\sum_{j,k\in V} y^i_{j,k} \cdot \ln(p(d_{j,k})) \tag{3.30}$$

Thus, the objective for maximizing the reliability can be written as

$$\text{Obj: maximize } \min_{i \in V_S} \sum_{j,k \in V} y^i_{j,k} \cdot \ln(p(d_{j,k})) \tag{3.31}$$

However, this objective will only maximize the minimum reliability of the network and will not maximize the reliability of all connections. Further, the reliability of a link is dependent on the number of connections that uses that link. Therefore, we will use an objective for every connection:

$$\text{Obj: maximize } \sum_{j,k \in V} y^i_{j,k} \cdot \ln(p(d_{j,k})); \ \forall i \in V_S \tag{3.32}$$

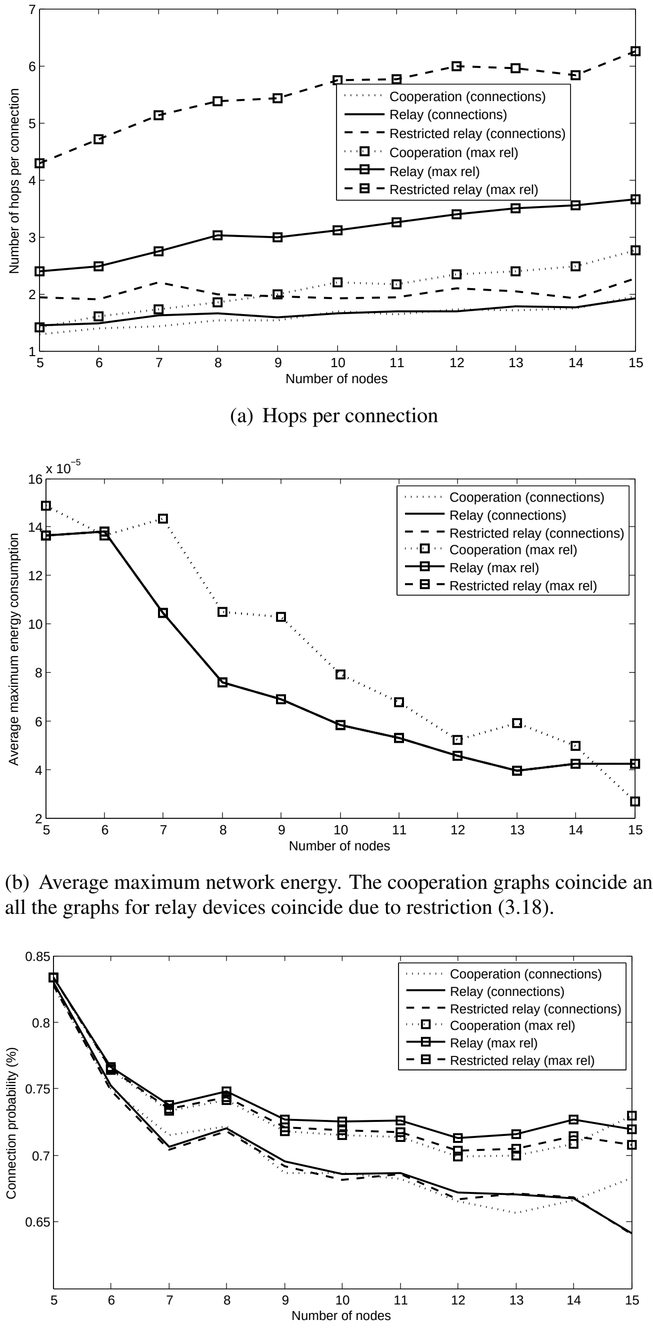
We still use an ILP consisting of 2 phases: first the maximum node energy is minimized using (3.17) and second the reliability is maximized by (3.32) for every connection while the energy consumption is restricted by (3.18).

The number of normal nodes is once again varied between 5 and 15 and the same number of relay devices as the number of nodes in the network is used. The comparison between objective (3.32) and the objective that minimizes the number of nodes (3.19) is shown in Figure 3.22 for the scenarios without and with the relay devices and the restricted number of relay devices. As expected, the number of hops increases when a more reliable path is chosen, Figure 3.22(a). The difference with and without relay devices is about 1 hop per connection. This means that every node sends its data to a relay device. The average maximum network energy is the same when the reliability is taken as objective as the maximum network energy is restricted by (3.18). All the graphs for relay devices coincide. In Figure 3.22(c) the resulting connection probability is shown. The reliability is almost the same with or without relay devices. This is because the relay devices are placed next to the normal sensor devices. When the reliability is used as objective, more reliable routes are chosen. Further, when the number of relay devices is restricted, the reliability drops a little bit.

3.6.3 Collision Probability

In the previous sections, it was assumed that no collisions occurred when packets where transmitted. This assumption was taken to orthogonalize the results from the properties of specific MAC-protocols. When the MAC-protocol is TDMA-based and the slots are well allocated, the results will remain the same. However, problems arise when the slots are not synchronized anymore. For example, consider the topology of Figure 3.20. When node A is sending to node B and node C is sending to node D, the packets will collide at node B when the duty cycle is not synchronized.

We can calculate the probability of collision when we know the number of neighbors and the activity level of the nodes. When the node has less neighbors,



(a) Hops per connection

(b) Average maximum network energy. The cooperation graphs coincide and all the graphs for relay devices coincide due to restriction (3.18).

(c) Reliability

Figure 3.22: ILP where the reliability is maximized. The restricted relaying is where the number of relaying nodes is minimized.

the number of nodes that can influence each other is limited. Consequently the interference will drop. We will estimate the probability of a collision at a node i:

$$\Pr[\text{collision i}] = 1 - \prod_{j \in NH(i)} (1 - \Pr[Act^j]) \tag{3.33}$$

The product is over all the neighbors of node i, $NH(i)$. $\Pr[Act_j]$ is the probability of activity of node j[1]. If we assume that the probability of activity at all nodes in the network is independent and identically distributed, then $\Pr[Act_j] = \Pr[Act]$ $\forall\, j \in\ NH(i)$. Consequently, equation (3.33) can be formulated as

$$\Pr[\text{collision i}] = 1 - (1 - \Pr[Act])^{NB} \tag{3.34}$$

where NB is the number of neighbors ($\#NH$). Plotting the probability, we get the following plot for a varying number of neighbors, Figure 3.23.

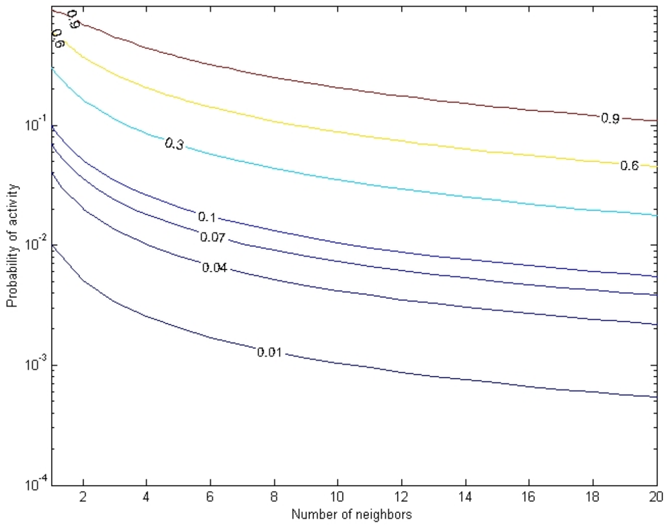


Figure 3.23: Probability of collision of node i for a varying number of neighbors. The X-axis shows a varying number of neighbors, the Y-axis a varying probability of activity. The lines on the graph show the places with a constant probability of collision.

We see that for a lower number of neighbors, the probability of collision will drop keeping the same probability of activity. Thus, by limiting the number of

[1] Notice that this model can take into account the difference between the communication range and the interference range of a radio. In regular wireless networks the interference range is about two times the communication range. In this study $NH(i)$ is defined as the set of nodes that can interfere with node i

neighbors, less interference will be experienced. The number of neighbors can be lowered by decreasing the transmit power $P_{s,dB}$. This will lead to a multi-hop network.

The number of neighbors will also influence the reliability and the energy efficiency. When a node has a lot of neighbors, the probability of a collision rises, leading to lower reliability and raising retransmissions. This will also lower the energy efficiency. When the nodes are grouped in a tree network and the tree is built in such a way that neighboring nodes reside in the same branch of the tree, collisions can be avoided by assigning slots. This problem will be the focus of Chapters 5 and 6.

3.7 Conclusion

In this section, we have investigated possible network topologies for a WBAN. Two important characteristics were considered: energy efficiency and reliability. For the former, we first defined the lifetime of the network as the time for the first node to die. We started with a very basic comparison between multi-hop and single-hop communication for a line topology and a tree network. It showed that multi-hop communication is better for nodes far away from the personal device, but not for nearer nodes. The effect of aggregation was taken into account and it was demonstrated that aggregation lowers the energy consumption, especially in a tree network. Then cooperation was introduced: sensor nodes cooperate in delivering the data to the personal device. As formal analysis, an ILP formulation was set up. Simulations have shown an increase of 24% on average of the network lifetime when the maximum energy per node is minimized. In a second step, the number of hops per connections was minimized. After minimization, this is about 1.5 hops per connection, indicating that cooperation indeed is beneficial for the lifetime of the network. Next, relay devices were introduced to lower the maximum network energy even further. And finally the concept of cooperation and the use of relay devices was combined and analyzed using an ILP. By letting nodes cooperate and by adding extra relay devices for a better load balancing, even more energy efficient network topologies can be obtained. For larger networks, a lifetime increase up to 30% was achieved. For future purposes, it could be interesting to investigate more comprehensively the impact of data aggregation on the lifetime of more general networks by adding it to the ILP.

The reliability was discussed briefly and the loss probability and probability of collision was investigated. We have compared the reliability between a single-hop and multi-hop architecture. Through an example with a line topology, it was clear that a multi-hop architecture is a more reliable choice. In order to evaluate this assumption, we have analyzed the reliability by adding it to the ILP formulation. It is concluded that the reliability can be increased with only a very limited impact on the network lifetime.

Overall, it can be concluded that a multi-hop architecture is the best choice for a WBAN and that one has to deal with a trade-off between energy efficiency and reliability. Adding relay devices is helpful for both the energy efficiency and reliability.

References

[1] T. Zasowski, F. Althaus, M. Stager, A. Wittneben, and G. Troster. *UWB for noninvasive wireless body area networks: channel measurements and results*. In Ultra Wideband Systems and Technologies, 2003 IEEE Conference on, pages 285–289, November 2003.

[2] V. Shankar, A. Natarajan, S. K. S. Gupta, and L. Schwiebert. *Energy-efficient protocols for wireless communication in biosensornetworks*. In Personal, Indoor and Mobile Radio Communications, 2001 12th IEEE International Symposium on, volume 1, San Diego, CA, USA, September 2001.

[3] P. J. Riu and K. R. Foster. *Heating of tissue by near-field exposure to a dipole: a modelanalysis*. IEEE Transactions on Biomedical Engineering, 46(8):911–917, August 1999.

[4] Q Tang, N. Tummala, S. K. S. Gupta, and L. Schwiebert. *Communication scheduling to minimize thermal effects of implanted biosensor networks in homogeneous tissue*. IEEE Transactions on Biomedical Engineering, 52(7):1285–1294, July 2005.

[5] J. Haapola, Z. Shelby, C. Pomalaza-Raez, and P. Mahonen. *Cross-layer energy analysis of multihop wireless sensor networks*. In Wireless Sensor Networks, 2005. Proceeedings of the Second European Workshop on, pages 33–44, January/February 2005.

[6] M. Bhardwaj, T. Garnett, and A. P. Chandrakasan. *Upper bounds on the lifetime of sensor networks*. In Communications, 2001. ICC 2001. IEEE International Conference on, volume 3, pages 785–790, Helsinki, Finland, 2001.

[7] E. Fasolo, M. Rossi, and M. Widmer, J.and Zorzi. *In-network aggregation techniques for wireless sensor networks: a survey*. Wireless Communications, IEEE [see also IEEE Personal Communications], 14(2):70–87, April 2007.

[8] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan. *An application-specific protocol architecture for wireless microsensor networks*. IEEE Transactions on Wireless Communications, 1(4):660–670, October 2002.

[9] Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin. *Directed diffusion: a scalable and robust communication paradigm for sensor networks*. In MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking, pages 56–67, New York, NY, USA, 2000. ACM.

[10] A. Natarajan, M. Motani, B. de Silva, K. Yap, and K. C. Chua. *Investigating network architectures for body sensor networks*. In HealthNet '07: Proceedings of the 1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments, pages 19–24, New York, NY, USA, 2007. ACM.

[11] D. S. J. De Couto, D. Aguayo, B. A. Chambers, and R. Morris. *Performance of multihop wireless networks: shortest path is not enough*. SIGCOMM Comput. Commun. Rev., 33(1):83–88, 2003.

[12] M. D. Yarvis, W. S. Conner, L. Krishnamurthy, J. Chhabra, B. Elliott, and A. Mainwaring. *Real-world experiences with an interactive ad hoc sensor network*. In Parallel Processing Workshops, 2002. Proceedings. International Conference on, pages 143–151, 2002.

[13] A. Woo, T. Tong, and D. Culler. *Taming the underlying challenges of reliable multihop routing in sensor networks*. In SenSys '03: Proceedings of the 1st international conference on Embedded networked sensor systems, pages 14–27, New York, NY, USA, 2003. ACM.

[14] J. N. Laneman, D. N. C. Tse, and G. W. Wornell. *Cooperative diversity in wireless networks: Efficient protocols and outage behavior*. IEEE Transactions on Information Theory, 50(12):3062–3080, December 2004.

[15] R. W. Thomas. *Cognitive Networks*. PhD thesis, Faculty of the Virginia Polytechnic Institute and State University, June 2007.

[16] M. Bhardwaj and A. P. Chandrakasan. *Bounding the lifetime of sensor networks via optimal role assignments*. In INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, volume 3, pages 1587–1596, 2002.

[17] H. Zhang and J. C. Hou. *On the upper bound of α-lifetime for large sensor networks*. ACM Trans. Sen. Netw., 1(2):272–300, 2005.

[18] S. C. Ergen and P. Varaiya. *On multi-hop routing for energy efficiency*. Communications Letters, IEEE, 9(10):880–881, October 2005.

[19] W. Heinzelman. *Application Specific Protocol Architectures for Wireless Networks*. PhD thesis, Massachusetts Institute of Technology, 2000.

[20] ILOG CPLEX [online] http://www.ilog.com/products/cplex.

[21] BEgrid [online] http://www.begrid.be.

[22] T. Himsoon, W. Pam Siriwongpairat, Z. Han, and K. J. Ray Liu. *Lifetime maximization via cooperative nodes and relay deployment in wireless networks*. IEEE Journal on Selected Areas in Communications, 25(2):306–317, February 2007.

[23] S. Biswas and R. Morris. *Opportunistic routing in multi-hop wireless networks*. SIGCOMM Computer Communications Review, 34(1):69–74, 2004.

4

MOFBAN:
a Network Architecture
for Wireless Body Area Networks

In the previous chapter, we have discussed suitable network topologies for a WBAN. A next step could be the development of a network protocol that uses these network topologies. However, as was stated in Chapter 2, a WBAN should support heterogeneity of the nodes. Depending on the specific application of the WBAN (i.e. medical monitoring, sport monitoring, gaming etc.) some features such as power control, data aggregation and so on will be required, others not. Further, some nodes will be more capable than others (e.g. more energy, more memory etc.). For each layer, an application developer has to determine which protocols are most suited for the purposes of the intended application, and has to go through the complicated process of combining them into an optimized protocol stack. In this chapter, we present a lightweight framework for a network architecture so as to allow for an easy and flexible adaptation of network protocols for a WBAN. This framework is called MOFBAN (MOdular Framework for Body Area Networks) and will allow for an easy integration of existing and newly developed protocols and optimizes the energy-efficiency by using its modular structure. It further provides a uniform interface for application and radio designers. We will first discuss existing cross-layer techniques in wireless networks and describe the different components (called modules) of the framework.

4.1 Introduction

4.1.1 Cross-layering in Wireless Networks

A myriad of protocols for sensor networks has been proposed in the last few years, taking on issues such as medium access, routing and data manipulation. Most of the early developed protocols are based on a strict layered structure [1, 2]. This way, a specific level needs no information about the inner workings of lower or higher levels. Functionalities at different layers can be altered without any impact on the other layers. This strict separation has proved to be a good solution for wired networks, allowing the protocol designers to focus only on a small subset of network functionality. However, it is not suitable for wireless networks [3, 4]. By using a cross-layer architecture, optimization can be done at several layers at once, it is possible to achieve a global optimization and conflicting optimizations between different layers can avoided. Hence, most recent protocols have opted for a more holistic view using a cross-layer approach.

Many cross-layer architectures have been proposed for ad hoc networks [3, 5, 6] and in lesser extent for sensor networks [7–9]. An overview of different cross-layer architectures is given in Fig. 4.1. Fig. 4.1(a) shows the traditional layered structure. In the approach of Fig. 4.1(b), new interfaces between the existing layers are defined. An overview of protocols adopting this method can be found in [10]. A major drawback of this approach is the introduction of many dependencies between the different layers. A more holistic view is used in Fig. 4.1(c). A shared database is used, accessible by all the layers. Retrieval of parameters incurs additional overhead, but the resulting architecture is more clear. Another class of cross-layer design is the complete merge of two or more layers of the protocol stack, e.g. the MAC-layer and the routing layer as seen in Fig. 4.1(d). The last approach, shown in Fig. 4.1(e), completely discards the layered structure. The required functionalities are implemented in different modules [11] or heaps [12] which interact and can be changed easily. This allows for a rich interaction between the building blocks of the protocol, but requires a total new approach which changes the very way protocols are organized.

4.1.2 Building an Architecture

The cross-layer approach forms a first step towards a more uniform architecture for sensor networks. In order to boost the development of sensor networks, all the protocols defined for sensor networks need to be glued together. This vision is also stated in [13] where they claim that the limiting factor in progress in sensor networks is the lack of an overall sensor network architecture. They propose an overall sensor network architecture (SNA) based on the sensor protocol (SP) [14], which forms a *narrow waist* that sits between the network layer and link layer. SP

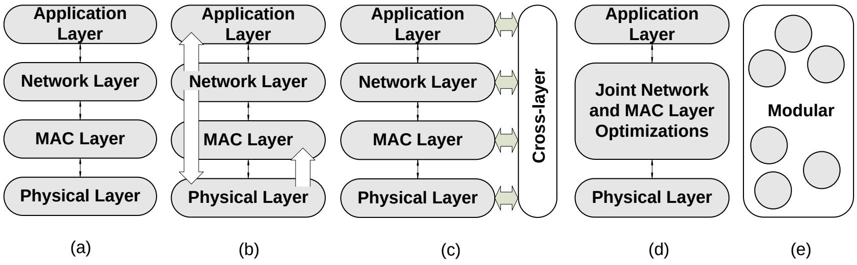


Figure 4.1: Overview of different cross-layer approaches. (a) Traditional layered structure (b) Passing of parameters across several layers (c) Holistic approach maintaining the different layers (d) Optimalizations spanning several layers (e) Modular approach.

is responsible for handling data transmission, data reception and neighbor management. However, SNA does not take into account the heterogeneity of the sensor nodes and has limited support for advanced network functionalities.

Another approach of a modular sensor network, is to make the sensor nodes themselves modular in hardware. This is done in [15] and in mPlatform [16], that both use a collection of stackable hardware modules that share a well defined common interface.

4.2 A Modular Approach

A long-term solution should be optimized to support heterogeneity and be adaptable to future advancements such as the creation of new physical carriers (e.g. the use of UWB-radios). Therefore, an architecture is proposed where the functionality is divided in modules that interact with each other. This modular approach has several advantages.

- Duplication of functionality can be avoided. Classic examples of duplicate functionalities are error correction and retransmission which are currently implemented in several layers of the protocol stack;

- Depending on the capabilities of the node, more modules (and thus network functionality) can be added. This way, heterogeneous networks can be supported;

- By allowing inter-modular parameter exchanges, cross-layer optimizations are possible. This results in much more energy-efficient protocols. Protocol information (such as neighbor tables) can be shared, resulting in better cooperation between protocols and less storage overhead;

- Through the replacement of modules, it is easy to adapt to changing network conditions and future developments.

4.2.1 Using a Modular Approach

When faced with an application scenario, application developers can choose which modules are required for a working solution (for example, whether or not mobility or QoS support is needed). The framework should also benefit protocol developers: the interaction of the new protocol with existing network solutions can easily be investigated by choosing different implementations of the interacting modules. Thus, the freedom when choosing the appropriate modules allows for detailed fine-tuning of the framework for specific applications and allows for advanced testing of different networking aspects. Thus, various applications and networks with diverging requirements can be designed on the same foundations.

Developing protocols in a modular way requires a new approach in designing network protocols. In a layered design, the interfaces between the protocols residing on different layers are well-defined, i.e. in the TCP/IP stack. This is no longer the case in a modular framework. One has to make sure the modules are called in the appropriate order and that the exchange of parameters between the modules is standardized.

4.2.2 Challenges when Designing Modular Protocols

Developing a cross-layer protocol is a promising but complex issue [10]. It is not an easy task to define the interactions between the layers, one can loose the compatibility with existing standards and optimization can lead to the creation of complex algorithms. In some cases, unintended cross-layer interactions can lead to an undesired deterioration of the global system performance [17].

Developing protocols in a modular way is a new paradigm and requires adjustments on the approach of designing network protocols. Some challenges when designing functional modules are the following [18]:

- The parameters to be exchanged by the different modules have to be determined. The exchanged parameters should have a substantial impact on the performance of the modules and the global system. The performance increase should at least compensate for the additional complexity introduced by the cross-modular interaction.

- There is a need to identify which parameters need to be optimized at design-time and which at run-time. When we are optimizing at design-time, optimal operation points are calculated off-line for various predefined operational conditions. These operation points are then used at run-time. This means that a look-up table is used which maps the optimal operational points for given operational settings. Generally, run-time solutions (aka real-time solutions) are more flexible and hence are better suited for dynamic networks, but these will require more complex nodes, leading to larger energy consumption. Depending on the complexity of the problem and the nature of the applied modules, real-time or off-line solutions may be adopted.

- The framework can be adaptive to changing network or application conditions. E.g. it could be beneficial to have more than one routing module and to activate one of them depending on the network/application environment.

- Due to the modular design, modules depend on the information of other modules. However, when replacing modules, it is possible that not all of the required information is available to the framework. When designing modules, one should keep in mind how the module will react when other modules do not supply some key parameters.

- When designing modules, careful consideration should be given to mutual dependencies. In particular, circular dependencies (where a steady-state can not be obtained) should be avoided. For example, when choosing the next hop node, the decisions of the routing module (shortest path) may interfere with decisions of the QoS module (reliable link) or the energy management module (hop with most remaining battery power). To prevent unceasing competition, a priority system or arbiter (taking into account the different preferences) should be developed. The interaction between different modules should thus be thoroughly examined.

4.3 A Modular Framework for WBANs

In this section, we present MOFBAN, a lightweight MOdular Framework for wireless Body Area Networks. The framework uses a modular design instead of the normal layered approach. This means that all the functionalities needed are implemented as software modules. The use of modules allows for a more flexible solution as some functionalities can be changed, added or removed more easily, simply by altering the corresponding module. These modules can be used and altered more easily compared to a general cross-layer approach where all the functionalities are implemented in one layer. In MOFBAN, the interaction between the different modules is handled by a controller module that is responsible for scheduling the appropriate function at the right time. The controller module further acts as a general access (an arbiter) to a storage space or data base for the parameters that can be used in the network.

An example of the MOFBAN framework can be found in Figure 4.2. The most important element of the framework is the middle part. It contains the differ-

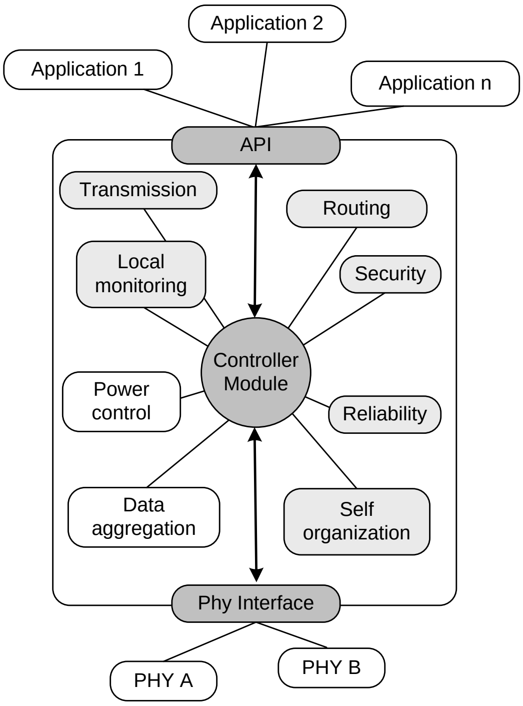


Figure 4.2: The modular framework for WBANs or MOFBAN

ent modules which implement the desired functionalities and holds the controller module. Further, two interfaces are provided. At the top, the application interface eases the use of the framework for application designers. The application layer is handled separately in order to facilitate the use of different applications (such as sensing, video,...) and the addition of extra services/applications. It is necessary to define a sort of generic interface between applications and the framework (see Section 4.5). At the bottom, the physical interface provides a uniform access to the physical layer. The framework has a simple architecture and the use of modules avoids the duplication of functionality, making the framework lightweight.

In a WBAN, different types of devices can be defined, see Section 2.1. The MOFBAN framework is implemented in all of these types:

- The *Sensor/Actuator nodes* are the regular nodes in the network. The specific use of the node is determined by the application designer who uses the application interface to activate the required functions. A node can be a ECG-sensor or an actuator. The question of which modules need to be used is determined by the application designer. The sensor nodes can have different implementations between themselves. Doing so, the network can support the different requirements of heterogeneity of the network.

- The *Relay devices* are merely used for relaying data, not for sensing. They do not have a need for interaction with applications as they do not run any.
- The *Personal device* acts as a gateway between the WBAN and other networks. Therefore, the personal device needs to support the normal IP-stack. It is an IP-capable device and it takes care of the conversion between the modular protocol stack and the layered OSI protocol stack. This device generally has two or more physical interfaces: one to connect with the WBAN and the other to connect with an external network. It further has a larger energy supply and more computation power.

4.4 Modules

The modules of MOFBAN take care of the networking functionality in general, such as routing, medium access, reliability, self organization,... Three types of modules can be distinguished: the *controller* module and the *mandatory* and *optional* modules. The mandatory modules are essential to have a proper working WBAN and implement the basic networking functionalities. When a more functional WBAN is needed, i.e. a WBAN supporting a certain level of reliability, security or power control, extra optional modules are added. It is up to the user of the network, i.e. the application designer, to define what is needed.

4.4.1 Controller Module

The controller module is the most important module and the heart of the modular framework. It is responsible for handling the correct functioning of the framework and handles all incoming requests from the physical and application layer. It consists of two major parts: a scheduler that calls the other modules at the appropriate time and a database that acts as a data-repository accessible by the other modules.

It is of utmost importance to keep the information between the different modules consistent. Therefore, the controller contains a database and it defines the interaction or communication process between the modules. If a module wants to access or pass information, this is communicated to the controller that retrieves the information from the database or stores the information in it. The controller module is responsible for solving and controlling interdependencies between different modules. A uniform data structure is used by all the modules. The controller module acts as a gateway to the database and controls the access. This is further discussed in Section 4.5.2.

The controller is further responsible for calling the appropriate module. This can occur packet-based or periodically. When a packet is received by the framework, the controller intercepts it. The data part is removed and temporally stored in the database. Based on the header information, the controller sets up a scheduling sequence. According to this sequence, the correct module is activated by the scheduler. When the module has finished, it passes the control back to the scheduler which activates the next module. The module is started in a new thread, so it does not block the working of the controller. The controller also holds several timers. Upon expiration of a timer, the appropriate module is activated.

4.4.2 Mandatory Modules

The mandatory modules are essential to have a proper working WBAN. In the following, an overview of the mandatory modules is given:

- The *Medium Access Module* regulates the transmitting of data on the medium and handles the channel access to the medium. The implementation in this required module is very basic: a simple CSMA with collision avoidance. The actual transmission happens in the physical layer.

- The *Routing Module* is responsible for setting up a path towards the personal device or other destinations. This can be done using a weight function, number of hops,.... The routing module provides the next hop to the transmission module via the controller module. It can use information about the network, e.g. from the QoS module or from the database in the controller. Different implementations can be made for this module. If a new protocol is used, a new routing module can be added easily. Even multiple routing modules can coexist in the same node. It is the responsibility of the controller to activate the correct routing module.

- The *Local Monitoring Module* monitors the network parameters such as the link quality, received signal strength, remaining battery power of the node, the number of neighbors,... Stated otherwise, this module retrieves the information of the physical layer and other layers that needs to be shared among the other modules. The information is stored in the controller's database.

- The *Self Organization Module* is responsible for the automatic set up of the network and for maintaining the network. It is activated by the controller module when the node starts up and at certain periods of time.

4.4.3 Optional Modules

The second type of modules are the optional ones. These modules are used to enhance the functionality of the network. Some examples to illustrate the realm of possibilities:

- The *Advanced Medium Access Module* introduces a more sophisticated channel access where slots are used. The mandatory module only implements contention-based medium access whereas this module implements contention-free medium access.

- The *ACK-Module* is used for sending acknowledgments. It stores a copy of the sent packet and a timer is started in the controller. If no ACK is received when the timer expires, the transmission module is activated by the scheduler. The module is also activated when a packet is received and it is required to send an ACK. This module mainly considers the reliability of one link.

- The *Security and Privacy Module* has as primary goal to authenticate devices in order to protect the Body Area Network from intruders. The second goal is to protect the privacy of sensitive information, i.e. medical data, by encrypting the data and/or encrypting the links between the authenticated devices. In order to obtain this, a private key can be exchanged.

- The *Power Control Module* acts on the transmission power of a node. This can be useful for limiting the number of neighbors and thus influencing the interference between nodes, or for lowering the power consumption and thus lengthen the lifetime of the node. This module not only considers the power control of each node individually, but can also look at the whole (or a large area) of the network. The module needs to work closely together with the routing module. The routing module can ask to this module to alter the transmission power (and as a consequence the network topology) if no appropriate route to the destination can be found.

- The *Reliability Module* can be regarded as an extended version of the transmission module. It adds error control, more advanced retransmission, priority queuing and so on. Stated otherwise, this module is responsible for providing end-to-end QoS and guaranteed delivery. If wanted, one can define multiple modules that each take care of a QoS aspect.

- The *Data Aggregation Module* can be considered as an extension of the routing protocol. The data received by other nodes is aggregated before sending them further to the personal device. Doing so, fewer transmissions are needed and the aggregated data can consist of fewer bits. Thus, the energy efficiency is increased. In this module, data aggregation is defined as putting the payload of several packets into one packet, without doing any processing on the data (i.e. concatenation).

- The *Local Data Processing* module adds the possibility to perform local data processing or in network processing such as taking the average of the received data.

- The *Mobility module* reacts upon the movement of the nodes. In a WBAN, only low mobility is required. The nodes itself will not be moving on the body, but mobility is caused by changes of the position of the body, e.g. when the person is running, moving his arm, etc. As the mobility is limited, an option would be to combine the mobility with the self organizing aspect. Mobility can mainly be considered as devices appearing and disappearing and this aspect can be handled by the self organizing mechanism. On the other hand, the mobility of the devices can have a periodic structure (e.g. when a person is moving his arms when running: the distance between the wrist and the shoulder changes in equal periods). This behavior can be exploited in order to facilitate the routing in the network. The use of these periodic movements can be considered as an extension to the mobility support and as non-essential. This module is an aid to the routing module or transmission module.

4.5 Communication

When developing a framework, it is important to define how the communication in the framework should take place. This communication can be either within or between the framework and the application or physical layer or between nodes themselves. In this section, we will briefly describe how the communication is handled in MOFBAN.

4.5.1 Communication within the Framework

The modules in the framework need to communicate with each other. Generally, this communication is nothing more than passing parameters to a module or invoking another one. In the framework, this is handled by the controller. The module sends the parameters to the controller which stores it in the controller's database. By using the controller as gatekeeper to the database, data conflicts between modules can be avoided. When the controller invokes a module, it passes the required parameters to the module. When the module has finished, it returns the new or altered parameters to the controller who puts them in the database.

The interaction between the framework and the application interface is provided by a sort of API, usable by the application designers. The API eases the application development as the designers do not need to be aware of the underlying network characteristics. It allows the designers to adjust settings such as the required level of security, the maximum desired delay and the bitrate needed by the application. The application interface makes an abstraction of the underlying protocols and selects the modules needed to meet the characteristics of the network/application. The API informs the controller module which modules need to be included in the framework.

The properties of the physical layer largely depend on the design of the hardware. In order to have a common interface usable by the framework, a mapping between the proprietary characteristics of the hardware and the more generic properties used by the framework is provided. Examples of such information are bit error rate (BER), path loss, received signal strength and so on. Every time a packet has been received, additional transmission information can be communicated to the framework through the interface. The controller module receives this information and stores it in the database. The PHY-interface is in essence a hardware abstraction layer. It provides some general characteristics to the framework that is unaware of the type of radio used.

4.5.2 Communication between Nodes

A last type of communication can be found between similar modules located at different nodes. For example, two routing modules that want to exchange routes with each other. As in regular protocols and frameworks, the packet header is used for this purpose. However, as the framework is modular, i.e. modules can be added, removed or changed, the packet header should be capable of coping with this modularity. This can be done by using a modular header structure where each module can add one or multiple extensions to the header. This extension consists of 2 fields: the parameter type and the parameter data. The parameter type indicates which parameter is sent and the parameter data contains the value. An example of the header of a modular packet is shown in Figure 4.3. When a packet needs to be sent, the controller retrieves the corresponding parameter values and the payload from the database and creates a new packet. When a packet is received by a node, the header is analyzed by the controller module. The controller puts the data of the header in the database, determines the scheduling sequence, starts the first one and passes the required arguments to the module.

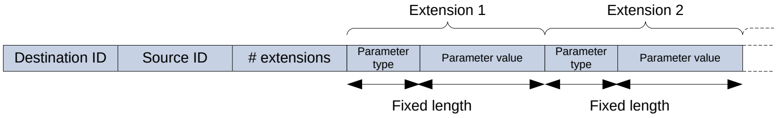


Figure 4.3: Header of a modular packet structure for MOFBAN.

The length of both the extension fields is fixed, which allows for an easier interpretation of the header. The header in a modular structure exists of the following fields: destination ID, source ID, number of extensions (say n), n extensions (consisting of parameter type and parameter value) and the payload.

4.6 Working of the Framework

An overview of the working of the framework can be seen in Figure 4.4. Assume that a packet is received by the Physical interface (1). This is forwarded to the controller (2). The controller now analyzes the header of the packet and extracts the parameters (3). These parameters and the payload are now stored in the database (4). When the controller has determined the scheduling sequence, it invokes the first module, passing along the required parameters (5). When this module has finished, it notifies the controller and sends back new or changed parameters. The controller now invokes the next module (6) and so on until the last module. This last module is for example a Transmission module that wants to send a new packet. The controller gets the required parameters and payload from the database (7) and constructs a new packet (8). This is sent to the Physical interface (9) that then transmits the packet (10).

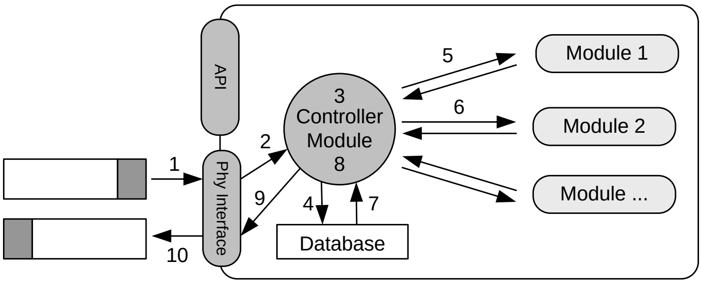


Figure 4.4: The working of MOFBAN. A packet is received by the controller who analyzes the header and puts the data in the database. The controller invokes the modules one after the other. At the end, the controller creates a new packet with the information from the database and transmits it.

4.7 Conclusion

The framework presented in this chapter serves as a preliminary proposal for a fully flexible modular architecture for WBANs. Various requirements considered in section 2.1 are covered, such as coping with heterogeneity, flexibility, ease-of-use, energy efficiency, ... However, several improvements can still be made.

It is designed in such a way that it is more transparent for the application designer. The functionality of the framework is easily adaptable and expandable.

This is made possible by the use of modules. Other resulting advantages are:

- Duplication of functionality is avoided and a simple structure is used. The framework can be regarded as being lightweight.
- Heterogeneity and QoS are well supported by using different implementations of the routing module or reliability module.
- Easy to add functionality by simply plugging in a new module in the framework.
- Quickly reconfigurable through the application interface.

Further, we have discussed the communication in the framework. The interaction between the modules is handled by the controller module and information is stored in a common database. Conflicts in the database are avoided by using the controller as gatekeeper. The controller module also takes care of the communication between modules, with the applications and physical layer en between the nodes. It further is responsible for activating the appropriate module at the correct time by using the scheduler. The controller module can be considered as the heart of MOFBAN.

In the communication part, we have introduced the concept of the modular header structure. This structure is beneficial for reducing the overhead between nodes as data can be shared better and no redundant or duplicate data is sent over the network.

The framework needs to be properly validated and the impact of the framework on the overall performance of the network should be thoroughly investigated. The overhead introduced by MOFBAN should be examined. We believe that this overhead will be minimal due to the avoidance of duplication and the use of the central database.

We are convinced that MOFBAN will proof to be a starting point for the development of new protocols for communication in a WBAN. It will ease the development of new applications and trigger the use of WBANs, and more generally WSNs[1]. In the next chapters, we will give an overview of existing WBAN network protocols and propose new ones. We will also show how these protocols can be put into the framework.

[1] The concept of a modular architecture for WSNs is part of ongoing research at IBCN. It will be further investigated and developed by Eli De Poorter

References

[1] I. F. Akyildiz, Weilian Su, Y. Sankarasubramaniam, and E. Cayirci. *A survey on sensor networks*. IEEE Communications Magazine, 40(8):102–114, August 2002.

[2] Ivan Stojmenovic, editor. *Handbook of Sensor Networks: algorithms and architectures*. Wiley-Interscience, December 2005.

[3] V. Srivastava and M. Motani. *Cross-layer design: a survey and the road ahead*. Communications Magazine, IEEE, 43(12):112–119, Dec. 2005.

[4] I. Chlamtac, M. Conti, and J.. Liu. *Mobile ad hoc networking: imperatives and challenges*. Ad Hoc Networks, 1(1):13–64, 2003.

[5] S. Toumpis and A. J. Goldsmith. *Performance, optimization, and cross-layer design of media access protocols for wireless ad hoc networks*. In Communications, 2003. ICC '03. IEEE International Conference on, volume 3, pages 2234–2240, May 2003.

[6] M. Conti, G. Maselli, G. Turi, and S. Giordano. *Cross-layering in mobile ad hoc network design*. Computer, 37(2):48–51, February 2004.

[7] P. Marron, A. Lachenmann, D. Minder, J. Hahner, K. Rothermel, and C. Becker. *Adaptation and cross-layer issues in sensor networks*. In Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004. Proceedings of the 2004, pages 623–628, December 2004.

[8] H. Kwon, T.H. Kim, S. Choi, and B. Gi Lee. *A Cross-Layer Strategy for Energy-Efficient Reliable Delivery in Wireless Sensor Networks*. IEEE Transactions on Wireless Communications, 5(12):3689–3699, December 2006.

[9] R. Madan, Shuguang Cui, S. Lall, and N. A. Goldsmith. *Cross-Layer Design for Lifetime Maximization in Interference-Limited Wireless Sensor Networks*. IEEE Transactions on Wireless Communications, 5(11):3142–3152, November 2006.

[10] T. Melodia, M. Vuran, and D. Pompil. *The State of the Art in Cross-Layer Design for Wireless Sensor Networks*. In EuroNGI Workshop on Wireless and Mobility, LNCS 3883, pages 78–92, July 2005.

[11] E. De Poorter, B. Latré, I.Moerman, and P. Demeester. *Universal Modular Framework for Sensor Networks*. In International Workshop on Theoretical and Algorithmic Aspects of Sensor and Ad-hoc Networks (WTASA'07), Miami, USA, June 2007.

[12] R. Braden, T. Faber, and M. Handley. *From protocol stack to protocol heap: role-based architecture*. SIGCOMM Comput. Commun. Rev., 33(1):17–22, 2003.

[13] D. Culler, P.l Dutta, C. Tien Ee, R. Fonseca, J. Hui, P. Levis, J. Polastre, S.Shenker, I. Stoica, G. Tolle, and G. Zhao. *Towards a sensor network architecture: lowering the waistline*. In HOTOS'05: Proceedings of the 10th conference on Hot Topics in Operating Systems, pages 24–24, Berkeley, CA, USA, 2005. USENIX Association.

[14] C. T. Ee, R. Fonseca, S. Kim, D. Moon, A. Tavakoli, D. Culler, S. Shenker, and I. Stoica. *A modular network layer for sensorsets*. In OSDI '06: Proceedings of the 7th symposium on Operating systems design and implementation, pages 249–262, Berkeley, CA, USA, 2006. USENIX Association.

[15] A. Y. Benbasat, S. J. Morris, and J. A. Paradiso. *A wireless modular sensor architecture and its application in on-shoe gait analysis*. In Sensors, 2003. Proceedings of IEEE, volume 2, pages 1086–1091, October 2003.

[16] D. Lymberopoulos, N. B. Priyantha, and F. Zhao. *mPlatform: a reconfigurable architecture and efficient data sharing mechanism for modular sensor nodes*. In IPSN '07: Proceedings of the 6th international conference on Information processing in sensor networks, pages 128–137, New York, NY, USA, 2007. ACM Press.

[17] V. Kawadia and P.R. Kumar. *A cautionary perspective on cross-layer design*. IEEE Wireless Communications, 12(1):3–11, February 2005.

[18] E. De Poorter, B. Latré, I. Moerman, and P. Demeester. *Sensor and Ad-Hoc Networks: Theoretical and Algorithmic Aspects*, volume 7 of *Lecture Notes Electrical Engineering*, chapter Universal Framework for Sensor Networks. Springer, 2008.

5
Network Protocols for Wireless Body Area Networks

This chapter deals with the network protocols for WBANs. First, the MAC-layer is considered. An overview of existing protocols for WBANs and WSNs in general is given. The throughput and delay of the IEEE 802.15.4-protocol is studied and its usefulness in WBANs is investigated. Next, existing routing strategies are studied. This information is then taken into account for the development of a first new cross layer network protocol for WBANs that handles both the medium access and the routing aspects: WASP or *Wireless Autonomous SPanning tree*. After a description, the protocol is analyzed and implementation results are given.

5.1 Medium Access Control

The MAC-protocol is responsible for coping with overhearing, idle listening and collisions in a wireless environment. It strongly influences the energy consumption of the nodes in the network and is therefore a critical factor in a WBAN. In the following, we give an overview of existing MAC-protocols. As the first dedicated protocols for WBANS were only recently introduced, we also discuss MAC-protocols in wireless sensor networks. In Section 1.1.2 we have seen that most current implementations of WBANs use IEEE 802.15.4 as enabling technology. To that end, the throughput and delay performance of IEEE 802.15.4 have been studied and its suitability for WBANs is discussed.

5.1.1 Related Work

As there are only few MAC-protocols for WBANs, we will start with the research in the area of wireless sensor networks. The two major categories of current MAC-protocols designed for wireless sensor networks are contention-based and schedule-based [1, 2]. For the former, CSMA/CA is a typical example, while TDMA is a typical scheme for the latter. The advantages of contention-based approaches are the simplicity, its infrastructure-free ad hoc feature and good adaptability to traffic fluctuation, especially for low load. Schedule-based approaches on the other hand are free of idle listening, overhearing and packet collisions because of the lack of medium competition, but require tight time synchronization.

The most commonly used technique for reducing energy consumption in contention-based protocols is controlling the power and duty cycle of the radio.

The Sensor-MAC (S-MAC) protocol [3] uses scheduling to coordinate sleeping among neighboring nodes to avoid idle listening. A node tries to synchronize with its neighboring nodes. If no neighbor is found, it will choose a schedule to start with. These synchronized neighbors form “synchronized islands”. Neighbors in reach of two or more islands have to synchronize with all islands’ schedules. An extension of the S-MAC is the T-MAC protocol [4], which allows the nodes to go back to sleep when no traffic is detected for a certain time. WiseMac [5] and B-MAC [6] use a preamble sampling technique. Nodes wake up for a short period after intervals with fixed length. The wake-up period is at least as long as the preamble so whenever a node wakes up and it hears a preamble it knows it has to stay awake to receive packets after the preamble. Doing so, the receivers significantly reduce the energy used for idle listening. To reduce the overhead associated with long preambles, a strobed sequence of short packets allowing for fast shutdown and response is used in [7]. An ultra-low duty cycle MAC that combines scheduling and channel polling is presented in [8]. All suffer from synchronization overhead and periodic exchange of sleeping schedules. D-MAC [9] can be summarized as an improved Slotted Aloha algorithm in which slots are assigned to the sets of nodes based on a data gathering tree. The principal aim is to achieve very low latency for converge cast communications, but still be energy efficient. However, the set up of the tree is not flexible and collisions may occur when nodes on the same level try to send data.

One of the few MAC-protocols for WBANs was proposed by Lamprinos *et al.* [10]. They use a master-slave architecture and, to avoid idle listening, all the slaves are locked in the Rx-slot of the master and go in standby at the same time. The main drawback of this protocol is that some slaves will have a low duty cycle whereas the nodes that are serviced later have a higher duty cycle. The protocol was implemented nor simulated. An adaptation of this protocol was used in [11]. This protocol divides time into frames in which only one node is allowed to transmit. The scheduling order is derived by applying the Earliest Deadline First algorithm. Omeni *et al.* [12] propose a MAC protocol for a star-networked WBAN that supports TDMA to reduce the probability of collision and idle listening. Each slave node is assigned a slot by the central node. When an alarm occurs at one of the nodes, the node can be assigned an extra slot for direct communication. The protocol has been evaluated on a Sensium platform. The H-MAC protocol [13] uses the human heartbeat rhythm information to perform time synchronization for TDMA. The bio-sensors can thus achieve time synchronization without having to turn on their radio. The algorithm is verified with real world data but assumes a certain buffer. The simulations do not show the energy gain and the protocol is designed for a star-topology WBAN only.

Most current implementations of WBANs use IEEE 802.15.4 [14] or ZigBee [15] as enabling technology. Some implementations use Bluetooth (IEEE 802.15.1) [16]. This was developed as a cable replacement and does not support (or only very limited) multi-hop communication. It has a complex protocol stack and a high energy consumption compared to IEEE 802.15.4. It is therefore not suited to be used in a WBAN. In the following section, we will discuss IEEE 802.15.4 and its appropriateness for communication in WBANs.

5.1.2 IEEE 802.15.4

The IEEE 802.15.4 standard [14] is designed as a low power and low data rate protocol offering high reliability. It defines the physical layer and the medium access layer and uses a beaconed and unbeaconed version. For the physical layer, 27 communication channels in three different frequency ranges are defined in the industrial scientific medical (ISM) band: 16 channels in the 2.4 GHz band, 10 channels at 915 MHz and 1 channel at 868 MHz. The 2.4 GHz band is available worldwide and operates at a raw data rate of 250 kbps. The channel of 868 MHz is specified for operation in Europe with a raw data rate of 20 kbps. For North America the 915 MHz band is used at a raw data rate of 40 kbps.

In [17] and [18] (added as Appendix A) we have analyzed both analytically and experimentally the maximum throughput and minimum delay of the unbeaconed or unslotted version of the protocol. The exact formula for the throughput and delay of a direct transmission between one sender and one receiver is given. This is done for the different frequency ranges and address structures used in IEEE 802.15.4. The analysis is limited to the unslotted version as this one experiences the lowest overhead. In section 5.1.2.1 we recapitulate the main results. For the full mathematical background and an overview of the IEEE 802.15.4 protocol, the reader is kindly referred to Appendix A.

5.1.2.1 Throughput and Delay

In the beaconless mode, a simple CSMA/CA protocol is used. When a device wishes to transmit data, the device waits for a random number of back off periods. Subsequently, it checks if the medium is idle. If so, the data is transmitted, if not, the device backs off once again. The standard defines three types of addressing: 64 bit, 16 bit and no addresses. The acknowledgments can be omitted and the packet size is limited to 127 bytes at MAC-level, headers included.

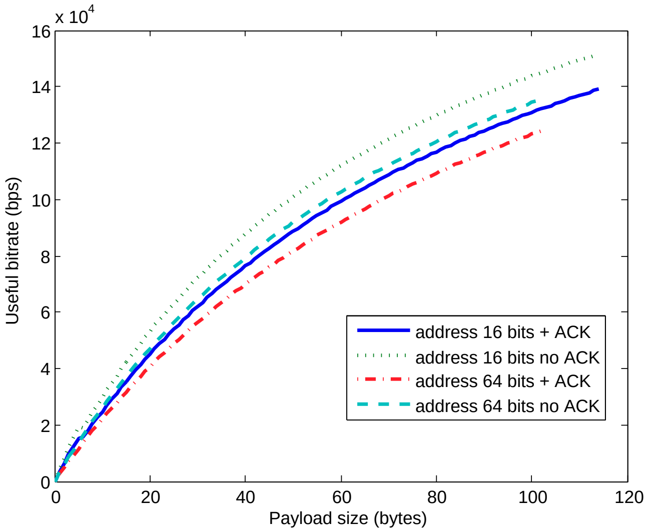


Figure 5.1: Useful bitrate for IEEE 802.15.4 in function of a varying payload size for the short and long address scheme, with and without ACK. The frequency is set to 2.4 GHz.

Figure 5.1 shows the maximum throughput for IEEE 802.15.4 in the 2.4 GHz region. The throughput is expressed as the useful number of bits, i.e. the data bits without the overhead or the MAC payload, that can be sent during 1 second. It is shown that the maximum throughput depends on the packet size. For smaller packets the throughput is significantly lower as the ratio of the overhead increases. The throughput is also lower when acknowledgments are used, as expected. When the addressing scheme of 16 bits is used, the overhead is smaller and the throughput is higher. The payload size can also be larger as the packet size is limited to 127 bytes including headers. In the 2.4 GHz band, a bandwidth efficiency of 64.9% is reached in optimal circumstances, i.e. when no addresses and no acknowledgements are used. If acknowledgements are used, an efficiency of merely 59.5% is obtained. Using the short address further lowers the maximum bit rate by about

4%. The worst result is an efficiency of only 49.8% which is reached when the long address is used with acknowledgements. The main reason for these low results is that the length of the MPDU is limited to 127 bytes. Similar conclusion can be made in the other frequency bands, see Table A.3.

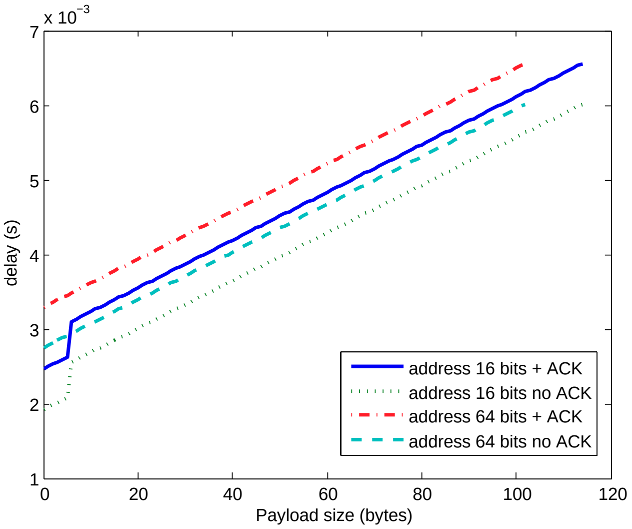


Figure 5.2: Minimum delay for IEEE 802.15.4 as a function of the payload size. The frequency is set to 2.4 GHz.

Figure 5.2 gives the minimum delay each packet experiences for varying packet sizes in the 2.4 GHz band. The minimum delay is calculated by sending a packet without any data bits immediately from 1 sender to 1 receiver. The propagation delay is not taken into account and no retransmissions are assumed. We immediately notice that the delay is a linear function of the number of payload bytes, as long as we assume a payload of more than 10 bytes. The transmitted frames are followed by an Inter Frame Space (IFS) in order to allow the MAC layer a finite amount of time to process data received from the PHY. Before starting the back off period, the device will wait one IFS. Long frames (MPDU is larger than 18 bytes) are followed by a Long IFS (LIFS) and short frames by a Short IFS (SIFS). The jump in the graph for the short address length is caused by the this mechanism. The same behavior is found for the other frequency bands.

The maximum delay is found by sending a full packet, i.e. the MPDU is set to the maximum of 127 bytes. The maximum delay is a little bit higher than 6 ms in the 2.4 GHz region when a full packet is sent, see Table A.4. This delay is acceptable for delay bound applications. The lower bands experience a significant higher delay, which is to be expected as the data rate is lower. In these frequency bands it is more important to look to the minimum delay, especially in the 868 MHz band.

Further we have also investigated the influence of the back off interval. A significant gain is found when the back off exponent is lowered. For more information, see Appendix A.

5.1.2.2 IEEE 802.15.4 on the Body?

In [19] the star network configuration of the IEEE 802.15.4 standard at 2.4 GHz was considered for a WBAN. The analysis considers quite extensively a very low data rate star network with 10 body implanted sensors transmitting data 1 to 40 times per hour. The analysis focuses on the effect of crystal tolerance, frame size and the usage of IEEE 802.15.4 Guaranteed Time Slots (GTS) on a node lifetime. The main consideration in this work was the long-term power consumption of devices. The results show that, even when properly configured, IEEE 802.15.4 provides a limited answer for medical sensor networking when configured in non-beacon mode with low data rate asymmetric traffic. Beacon mode may also be used, but with more severe restrictions on data rate and crystal tolerance.

Another adaptation is BSN-MAC [20]. The coordinator controls the communication by varying the superframe structure of IEEE 802.15.4. This divides the time axis in a contention-free and contention-based period. The sensors provide real-time feedback to a BSN coordinator with application-specific and sensor-specific information. Hence, based on the feedback the BSN coordinator can make dynamic adjustments for the length of the contention-free and contention-based period to achieve better performance in energy efficiency and latency.

Both [21] and [22] come to the conclusion that although 802.15.4 can provide QoS, the technology is not scalable in terms of power consumption and can not be used as a single solution for all WBAN applications. This view is shared by IEEE 802.15.6 that aims to develop a proper MAC for WBANs [23].

As such, it can be concluded that IEEE 802.15.4 is not the best solution for supporting communication in WBANs. Although it can be used for a quick (and easy) implementation, the results are rather poor. IEEE 802.15.4 was not designed to support WBANs. Specialized MAC protocols are needed. As a consequence, we will not use radios that implement the IEEE 802.15.4 PHY-layer, but more general low-power transceivers such as the Nordic nRF2404 transceiver (see Section 2.3.3.2).

5.2 Routing

A lot of research has already been performed in the area of wireless sensor networks. An overview can be found in [24]. In the following, an overview of existing routing strategies for WBANs is given.

When considering wireless transmission around and on the body, important issues are radiation absorption and heating effects on the human body. To reduce tissue heating the radio's transmission power can be limited or traffic control algorithms can be used. In [25] rate control is used to reduce the bioeffects in a single-hop network. Another possibility is a protocol that balances the communication over the sensor nodes. An example is the Thermal Aware Routing Algorithm (TARA) that routes data away from high temperature areas (hot spots) [26]. Packets are withdrawn from heated zones and rerouted through alternate paths. TARA suffers from low network lifetime, a high ratio of dropped packets and does not take reliability into account. An improvement of TARA is Least Temperature Routing (LTR) [27] that reduces unnecessary hops and loops by maintaining a list in the packet with the recently visited nodes. A combination of LTR and shortest path routing is Least Total Route Temperature (LTRT) [28]. The nodes' temperatures are converted into graph weights and minimum temperature routes are obtained. A better energy efficiency and a lower temperature rise is obtained, but the protocol has as main disadvantage that a node needs to know the temperature of all nodes in the network. The overhead of obtaining this data was not investigated.

“Anybody” [29] is a data gathering protocol that uses clustering to reduce the number of direct transmissions to the remote base station. It is based on LEACH [30] that randomly selects a cluster head at regular time intervals in order to spread the energy dissipation. The cluster head aggregates all the data and sends it to the base station. LEACH assumes that all the nodes are within sending range of the base station. Anybody solves this problem by changing the cluster head selection and constructing a backbone network of the cluster heads. The energy efficiency is not thoroughly investigated and reliability is not considered. Another improvement of LEACH is Hybrid Indirect Transmissions (HIT) [31], which combines clustering with forming chains. Doing so, the energy efficiency is improved. Reliability, however, is not considered.

Ruzelli et al. propose a cross-layer energy efficient multi-hop protocol built on IEEE 802.15.4 [32]. The network is divided into timezones where each timezone takes turn in the transmission. The nodes in the farthest timezone start the transmission. In the next slot, the farthest but one sends it data and so on until the sink is reached. The protocol almost doubles the lifetime compared to regular IEEE 802.15.4. The protocol was developed for regular sensor networks, but the authors claim its usefulness for WBANs.

This overview clearly shows that routing protocols for WBANs are an emerging area of research. The protocols described above were only developed in the last two years. A combination of MAC-layer and routing layer in one protocol has not been considered yet. In the following, we will present a cross layer protocol, WASP.

5.3 Wireless Autonomous Spanning Tree

In this section we present a slotted multi-hop approach to medium access control and routing in wireless Body Area Networks, the *Wireless Autonomous Spanning tree Protocol* or WASP. It uses crosslayer techniques to achieve efficient distributed coordination of the separated wireless links. This protocol incorporates slotted medium access and the automatic setup of the different routes. To that end, a spanning tree is set up automatically with the personal device as root. This spanning tree is subsequently used to route the data toward the personal device (or sink) and to assign the different slots in a distributed manner. Thus, in WASP, medium access control and routing are handled by the same spanning tree and therefore a higher throughput and lower energy consumption can be achieved. The WASP-protocol addresses the issue of intra-body communication. In the following, the term *nodes* refers to the sensor and actuator nodes.

5.3.1 Protocol Description

WASP is a slotted and distributed protocol that uses a spanning tree for medium access coordination and traffic routing. Each node tells its children in which slot they can send their data by using a special message: a *WASP-scheme*. This WASP-scheme is unique for every node and constructed in the node sending the scheme. A distinguishing property of WASP is the dual usage of the WASP-schemes by exploiting the broadcast nature of wireless links. The schemes are simultaneously used to control the traffic of the node's children and to request more resources from its parent for these children. This minimizes the coordination overhead as each scheme is used by both the parent as the children of the sending node.

All the information the node has to know in order to generate this scheme can be obtained by listening to the WASP-schemes coming from its parent node (i.e. one level up in the tree) and from its children (i.e. one level lower in the tree). Consequently, the division of the time slots is done in a distributed manner. No central management is needed.

Hereafter an example is given of the operation of the protocol in a steady-state situation. In this section we assume the spanning tree has already been constructed and is used to propagate the data to the personal device. The time-axis is divided in different cycles (further referred to as WASP-cycles) that consist of a fixed number of time slots. In a WASP-cycle, each node is allowed to send its data and/or to forward data received in the previous cycle to the next node. At the beginning of each cycle, the sink sends its WASP-scheme to its children. This WASP-scheme informs the sink's children when they can send their information. These children respond to the scheme by sending out their own WASP-scheme in their designated time slots. These WASP-schemes are based on the WASP-scheme of the sink and on the requirements of their children. Thus each node right below the sink calculates its own WASP-scheme and sends it to its own children which form the second level. On their turn, these nodes send out the WASP-scheme and so on. Doing so, each node will know when it can send its data without the need for a device that centrally calculates the distribution of the time slots. A more elaborate example will be given in section 5.3.3.2.

5.3.1.1 WASP-schemes

Each node sends a WASP-scheme to its children to inform them when they are allowed to use the wireless link. In this section, we will describe what elements can be found in such a WASP-scheme. In order to focus our thoughts, the simple example network of Figure 5.3 will be used. The full lines between the nodes indicate the tree structure of the network. All communication is wireless and the tree is constructed in such a way that the nodes only can hear their parent, their siblings (i.e. children of the same parent, dotted lines on the Figure) and their children. Thus there is no interference between adjacent branches of the tree. This assumption simplifies the protocol explanation. However, let it be clear that it is not necessary for two siblings to hear each other. We assume that in this example each node only has one data packet to send to the sink per cycle. Of course, in general, nodes will be allowed to send multiple data packets per cycle.

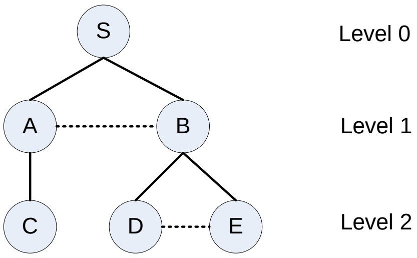


Figure 5.3: Example network. The full lines indicate the tree structure, the dotted lines the nodes that can reach each other.

The WASP-scheme of the personal device (referred to as node S)[1] is as follows:

$$\overbrace{S}^{1}\underbrace{AB}_{2}\overbrace{.^{3}}^{3}\underbrace{ABB}_{4}\overbrace{X}^{5}-\underbrace{11101}_{6} \tag{5.1}$$

where

[1] The personal device is often referred to as the sink. In the following, both terms will be used.

1 = The address of the sending node;
2 = Every child of the sink is designated one slot. In that slot, the node sends its WASP-scheme followed by data;
3 = A silent period (SP) of 3 slots. In this period, the child nodes receive data from their children that they need to forward to the sink;
4 = The children forward the received data to the sink;
5 = Contention slot;
6 = ACK-sequence.

First, each of the child nodes of the sink is awarded 1 slot, part 2 of scheme (5.1). In this slot, the child nodes will send their WASP-scheme and data. This WASP-scheme will be heard by the sink and by the children of the node. Consequently, the sink is informed of its child node's traffic requirements and is capable of calculating the duration of the silent period for the next cycle.

During the silent period, part 3 of scheme (5.1), the sink will not be sending or receiving any messages as this period is used by the sink's children to receive data or messages from their children. Doing so, interference between different levels will be limited.

Following the silent period (SP), a sequence is given in which the children can forward the received data to the sink, part 4 in (5.1). After that, a contention slot is inserted in order to allow new nodes to join the network, part 5 in (5.1). More information about the use of contention slots is to be found in section 5.3.1.3. The scheme is completed by the piggybacked acknowledgments, represented by simple bits, part 6 in (5.1). This will be explained in section 5.3.1.6.

A slightly different WASP-scheme is used by the other nodes, for example node A:

$$\overbrace{A}^{1}\overbrace{.^{1}}^{3}\underbrace{C}_{4}\overbrace{X}^{5}-\underbrace{1}_{6} \tag{5.2}$$

A first difference is the duration of the silent period. Whereas with the sink the silent period was used to allow the nodes of level 2 to send their data to level 1, the silent period is now used to indicate the nodes of level 2 when they should remain silent in order to not interfere with the ongoing communication on the higher level. In other words, in the silent period of the sink, the sink will be silent, and in the silent period of the other nodes, their child nodes will be silent. This reasoning brings us to the second difference with (5.1): the omission of part 2. The nodes below level 1 are not allowed to send their data immediately as all siblings of their parent node need to send their WASP-scheme first in order to reduce delay.

In the other levels (level three and lower), the nodes can start sending when the transmission of the parent node has ended, but the same principle holds: the silent period indicates when the child nodes should remain silent (and consequently can turn off their radio). A more elaborate example with a larger number of levels can

Slot	**Level 0**	**Level 1**	**Level 2**
0	*S*: WS (SAB.^{3}ABBX-10011)		
1		*A*: SD + WS (A.^{1}CX-1)	
2		*B*: SD + WS (BDEX-11)	
3			*C*: SD + WS (C.^{1}X-)
			D: SD + WS (D.^{2}X-)
4		*A*: CS	*E*: SD + WS (E.^{1}X-)
5		*B*: CS	*C*: CS
6		*A*: FD	*D*: CS
			E: CS
7		*B*: FD	
8		*B*: FD	
9	*S*: CS		

Abbreviations	WS	Sending WASP scheme (scheme between brackets)
	SD	Sending data
	CS	Contention slot
	FD	Forward data

Table 5.1: Steady-state WASP cycle for the example network of Figure 5.3. Notice that in slot 3 two nodes can send simultaneously. In slots 4, 5 and 6 some nodes have planned their collision slot simultaneously.

be found in section 5.3.3.2.

An overview of the entire WASP-cycle of this example can be found in Table 5.1. In each timeslot, the action taken is shown. In timeslot 0, the sink S, at level zero, sends its WASP-scheme. In the next timeslot, sensor A sends its WASP-scheme and its data. Further, in timeslot 3, we can see that node C and D can send simultaneously.

5.3.1.2 Calculation of the Duration of the Silent Period

The duration of the silent period for the sink SP_S equals the number of slots needed to send the data from level two to level one. Thus, it can be calculated as follows:

$$SP_S = \max_{\forall i \in Ch_S} \left(\sum_{j \in Ch_i} TS_j \right) + 1 \qquad (5.3)$$

where Ch_i are the children of node i, TS_j is the number of slots needed by node j to send the data to his parent, including its own data. V_i denotes the nodes in the tree below node i, node i not included. The maximum function is needed to indicate that the nodes on level 2 belonging to another branch of the tree (i.e. nodes with a different parent) can send simultaneously. One extra slot is inserted to account for the presence of a contention slot.

The length of the silent period of the other nodes is calculated based on the parent's WASP-scheme and is consequently different below level 1:

$$SP_{level_1} = \text{slot \# start silent period - slot \# node's first occurrence - 1} \tag{5.4}$$

$$SP_{level_i > 1} = \text{slot \# contention slot - slot \# node's first occurrence} \tag{5.5}$$

For example, the length of the silent period of node A equals the slot number of the start of the silent period (= 3) minus the slot number of the first occurrence of node A in the WASP-scheme of the sink (= 1) which gives a silent period of 1 slot. The length of the silent period of node C is the slot number of the contention slot (= 4) minus the slot number of the first occurrence of node C (= 3) which equals 1 slot. Indeed, a silent period of one slot is necessary because node A still has a contention slot.

5.3.1.3 Contention Slots

A contention slot allows nodes to join the network by sending a JOIN-REQUEST-message. It is possible to omit this slot for a few consecutive WASP-cycles to increase the maximum throughput. However, in order to keep the connection time reasonably low, at least one contention period per n milliseconds is required. Moreover, to support a notion of mobility a high frequency of contention slots is required. As it is possible that more than one node sends during these contention slots, a random delay is inserted before each node's transmission. Doing so, the probability of collisions during these slots is decreased. There is no special mechanism that detects collisions. If a node that wants to join the network does not hear its address in the next WASP-scheme of its parent, it assumes that a collision has occurred and sends a new JOIN-REQUEST in the next contention slot.

5.3.1.4 Building and Maintaining the Spanning Tree

The following steps are taken to let nodes join the network:

1. The new node scans the wireless medium for a certain time and picks up the WASP-scheme of the nodes in range.

2. Based on these received WASP-schemes, the new node determines which node will be the best parent, based upon the node's requirements (low delay, reliability, a weighted average, . . .).

3. During the contention slot of the chosen parent, the new node sends a JOIN-REQUEST.

4. The parent receives this request and adds the new node in the next WASP-scheme.

If for a certain time a child node does not receive a WASP-scheme, a new parent is chosen using the same procedure. Each node should periodically check which nodes are in its neighborhood and choose a new parent if a better one is available. The node does not have to notify its parent. If a parent doesn't receive the WASP-scheme of a child for n consecutive times, it assumes that the child is no longer part of the network and doesn't include it in the WASP-scheme anymore.

These basic tree maintenance mechanisms support mobility as nodes can (re-)join the tree. When a node is out of reach of its parent, a new parent is found using this mechanism.

5.3.1.5 Routing and Addressing

In order to facilitate the forwarding to the different nodes, the tree structure of the network can be used for addressing. However this poses problems as nodes might change attachment points in the tree. A more general solution is manually asking the sink for a unique address. The sink receives traffic from all sensors so it knows the addresses in use. A WASP address is composed of 6 bits. The sink gets address 000000, the first node will get 000001 and so on. Using 6 bits the number of nodes is limited to 64, which is reasonable due to the nature of a WBAN (see section 5.3.2.4). This addressing mechanism is not scalable for larger networks, but is useful for the small number of nodes in a WBAN.

In a WASP-scheme, 1 byte describes each slot. The first bit denotes the slot type. 0 stands for a regular slot where data is sent and 1 for a special slot. In case of a data slot the second bit denotes the traffic direction: 0 for regular traffic to the sink and 1 for traffic in the other direction or more general traffic that requires routing. The other 6 bits define the actual address of the node that is permitted to send in the slot. In case of a special slot, the 7 remaining bits are used to define the length of the silent period. E.g. 1.0000011 denotes a silent period of length 3. If all the bits are set to 1, the slot is a contention slot instead of a silent period.

Most traffic in a WBAN flows from the nodes to the sink. Traffic in the other direction can be supported by setting the second bit in the WASP-scheme. A node will add the source address with that bit set to the WASP scheme. Each child then decides whether it is on the path to the destination. If so it turns on its radio to receive the packet. Routing can be done by using a technique similar to learning bridges. Nodes record the addresses in traffic passing by and route packets from the sink to the nodes using that information.

5.3.1.6 Acknowledgments

At the end of each WASP-scheme, an ACK-sequence for the previous WASP-cycle is sent. It contains a bit for each slot in which data was sent in the previous cycle. A 0 denotes that the packet was not received correctly, a 1 denotes success. The node that receives the WASP-scheme, say node i, will check the ACK-sequence. If the position of a 0 corresponds to one of the slots where node i sent data, the node will resend that data in this cycle. Its parent will already include an extra slot for node i.

This acknowledgment scheme is quite weak but it fulfills our need. Due to the absence of contention, data loss will be limited to interference problems.

5.3.1.7 Heterogeneous Data Rates

In the example above, each node was only assigned one data packet or slot per cycle. In more general networks, some nodes will require a larger data rate then others. In such cases, these nodes will be assigned multiple slots per cycle. The desired number of slots can be requested when joining the network. Sometimes the wanted data rate of a node can change over time, e.g. when more accurate measurements are desired over a certain period of time. The node will ask for more slots in its WASP-scheme. This will be noticed by the parent node that will assign more slots in its next WASP-scheme. WASP also supports low duty nodes. Nodes that have no data to send during a certain time, can disconnect from the tree and rejoin the network when necessary by sending a JOIN-REQUEST to the parent. If desired, the node can remain connected and forward data from other nodes. Doing so, the node cooperates in the network and spreads the load of the network. If the node has data to send, an additional slot is requested. This mechanism also allows the addition of relay devices in the network.

5.3.1.8 Synchronization

The nodes in the tree need to be synchronized in order to avoid shifting of the start of slots between nodes. However the recurring cycles allow for resynchronization at the beginning of each cycle. Each node should wake up some milliseconds before the start of a slot to avoid these issues. If necessary, a more stringent synchronization can be obtained by letting the personal device send a pulse at the beginning of each cycle. This pulse should be more powerful than normal radio traffic so it is received by all nodes.

5.3.2 Performance Analysis

In this section, we present the performance analysis of our proposed protocol. We will address performance issues such as minimum delay, maximum throughput and sleeping time of the nodes.

5.3.2.1 Maximum Overall Throughput Efficiency

The maximum overall throughput efficiency of the protocol can be seen as the percentage of the useful traffic (or good put) that can be sent over the network. As we are working in a multi hop environment, the maximum throughput efficiency will mostly depend on the number of nodes in the network and the number of levels in the tree. Indeed, the maximum amount of data that can be sent to the sink per WASP-cycle depends on the number of nodes in the first level and the length of the silent period. As mentioned in section 5.3.1.1, the silent period of the sink allows the nodes of level one to receive their data. This means that if we can lower the duration of the silent period, the maximum throughput will rise. Thus by minimizing the number of children of the nodes of level one, the maximum throughput efficiency can be improved. Generally, the throughput efficiency can be found as

$$TPE \;=\; \frac{\text{\# slots where data is sent to sink}}{\text{length of a WASP-cycle}} \tag{5.6}$$

where the length of the WASP-cycle is expressed in slots and depends on the length of the silent period.

In the example of Figure 5.3, we see that per WASP-cycle 5 packets can be sent to the sink and the total length is 10 timeslots. Thus, we have a maximum throughput efficiency of $\frac{5}{10}$ or 50%. This seems to be a low number, but we have to keep in mind that we are working in a multi hop environment and that the medium is shared between multiple nodes.

The following formula determines the length of a WASP-cycle T_{WC}:

$$\begin{aligned} T_{WC} = &\text{\# data children } L_1 \;+\; SP_S \;+ \\ &\text{forwarding data of } L_2 \text{ from } L_1 \text{ to } L_0 \;+\; 2 \end{aligned} \tag{5.7}$$

where L_i represents level i. The two extra time slots added at the end are used for the transmissions of the sink's WASP-scheme and the contention slot at the end.

The duration of the forwarding period equals the number of timeslots needed to send the data of each node. Using (5.3) we can rewrite this formula as

$$T_{WC} \;=\; \sum_{i\,\in\,Ch_S} TS_i \;+\; \max_{\forall i\,\in\,Ch_S} \left(\sum_{j\,\in Ch_i} TS_j \right) \;+\; 3 \tag{5.8}$$

In order to simplify the formula and to make it more intuitive, we introduce δ_i. This denotes the number of packets node i generates in one cycle. Using this notation, we can write:

$$TS_i \;=\; \sum_{j\,\in\,V_i} \delta_j \;+\; \delta_i \tag{5.9}$$

Thus, (5.6) can be reformulated as

$$TPE = \frac{\sum_{i \in V_S} \delta_i}{\sum_{i \in V_S} \delta_i + \max_{\forall i \in Ch_S} \left(\sum_{j \in V_i} \delta_j \right) + 3} \tag{5.10}$$

and highly depends on the number of nodes and the structure of the tree. V_S refers to all the nodes in the network except the sink. In order to evaluate this formula, we assume that each node only has 1 packet to send per WASP-cycle. This means that the numerator of (5.10) equals the number of nodes in the network without the sink ($N - 1$). Further, we limit the maximum number of children a node can have, referred to as ζ. When ζ = 1, all nodes are in different levels and ordered in a line topology and when ζ = N - 1 all nodes can communicate directly with the sink. Of course, ζ can not be higher than N - 1. The second term in the denominator of (5.10) now needs to be written in terms of ζ. For simplicity, we assume a ζ-balanced tree, i.e. each node has exactly ζ children. The second term is the maximum number of nodes below a child of the sink. As the tree is regular by definition, the number of nodes is equally distributed between the children of the sink. Thus, we get

$$\max_{\forall i \in Ch_S} \sum_{j \in V_i} \delta_j = \left\lceil \frac{(N - 1) - \zeta}{\zeta} \right\rceil \tag{5.11}$$

where the right hand of the equation is rounded upwards.

Figure 5.4 shows the throughput efficiency for varying N and ζ. For 50 nodes, an efficiency of 94% is achieved for ζ = 1 and 49% for ζ = 49. When the size of the network is smaller, the efficiency is lower because of the use of the contention slots which is independent of the size of the network. When ζ is lower, the silent period is longer as more data needs to be forwarded in the tree. Thus, in order to increase the throughput efficiency, the silent period should not be too long. From the graph it can be concluded that ζ should be 5 or higher. The more nodes the network has, the higher the throughput efficiency. This is due to the fact that as more slots are used for sending data, the length of the WASP-cycle increases, see (5.8). The fixed part of the WASP-cycle (i.e. a slot for the sink and 2 contention slots) however remains the same. This will lead to a higher throughput efficiency.

5.3.2.2 Delay Limits

The experienced delay depends on the number of levels present in the network. Indeed, a node can only send his data up one level during each WASP-cycle. The only exception is to be found at level 1, where the sink's children can first receive the data from their children and then forward the data.

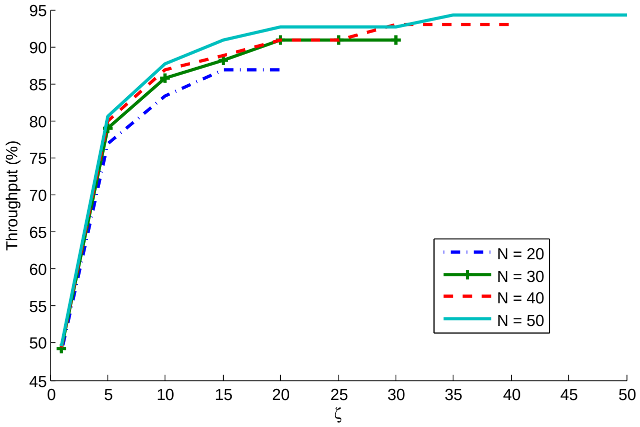


Figure 5.4: Throughput efficiency of WASP for varying number of nodes and changing tree topology with limited number of children per node.

We can define an upper and lower bound for a node i:

$$\text{Lower bound} = \max\left((level\ node_i - 2) \cdot T_{WC}, 1\right) \tag{5.12}$$

$$\text{Upper bound} = \max\left((level\ node_i - 1) \cdot T_{WC}, 1\right) \tag{5.13}$$

The maximum function is needed as the delay can not be lower than 1 slot.

The maximum delay over the whole network can be expressed as follows, assuming that the network has at least 2 levels:

$$\text{maximum delay} = \left(\max_{\forall i \in V}(level\ node_i) - 1\right) \cdot T_{WC}. \tag{5.14}$$

Summarizing, if we want a high throughput, we should minimize the length of the silent period and for a low delay minimize the number of levels. These two conditions do not contradict, therefore a high throughput can be achieved while preserving the low delay.

5.3.2.3 Sleep Ratio

When the nodes have heard the WASP-scheme of their parent, they can go into a sleep modus in the slots where they are not involved in the communication. This allows for energy saving. For example, the Nordic transceiver has two power saving modes: a standby mode consuming 12 μA and a power-off mode consuming 1 μA [33]. Even in standby mode, the power consumption is more than thousand times lower compared to the Rx or Tx mode. An important difference between these two modes is the switching time: switching from the standby mode to the normal mode takes less than 200 μs, from power-off mode to normal mode about 3 ms. This switching time is important for maintaining the synchronization. Hence, if a node can go to sleep mode for a longer period of time, the power-off mode is the most interesting choice. In the following, we assume that the radio does not consume any energy when it is switched off. This is reasonable as the power consumption in either the standby mode or the power-off mode are almost negligible and we are only considering the sleep ratio in one cycle.

In the example scheme of Figure 5.3, the sink can turn its radio off in the silent period as it will not receive data from its child nodes. Thus, the radio can be turned off 3 slots. Node A can also turn its radio off in its silent period, when it knows that its siblings are sending and when none of its children is allowed to send data. So, node A can sleep 5 slots.

The sleep ratio ρ_i of node i is defined as:

$$\rho_i \;=\; \frac{T_{WC} - T_{on,i}}{T_{WC}} \tag{5.15}$$

The following formula can be used to calculate the number of time slots in which node i has to operate its radio $T_{on,i}$:

$$T_{on,i} \;=\; (\sum_{j \in V_i} \delta_j + 1) + \delta_i + \sum_{j \in V_i} \delta_j + 1. \tag{5.16}$$

The first term refers to receiving the data from its lower layers (including the contention slot), the second and third term bring the sending of the data into account and in the last term, the node is listening to the scheme of its parent. This formula gives the upper bound of the number of slots a node can sleep. Indeed, if a node perfectly knows when a slot starts, it could turn on its radio at the beginning of each slot for a very short time. If we divide the formula by the duration of a cycle, we get the time ratio the node can sleep.

In a ζ-balanced tree the most burdened nodes are the nodes right below the sink. If we analyze the formula further assuming a regular ζ-balanced tree and using (5.11) for the node with the most children, the maximum time ratio for the nodes on the level below the sink can be written as

$$\rho \;=\; \frac{N - \left\lceil \frac{N-1-\zeta}{\zeta} \right\rceil - 1}{N + \left\lceil \frac{N-1-\zeta}{\zeta} \right\rceil + 2} \tag{5.17}$$

Figure 5.5 shows the sleep ratio for varying N and ζ. There is no big difference between large and small networks. When ζ is lower, the sleep ratio ρ is lower as the node will have more nodes below it (more nodes in V_i) and will have to relay more data. This can especially be seen for a ζ = 1 where the sink has only one child that has to forward all the data and the sleep ratio is only 5%. It can be concluded that, also for the energy efficiency, ζ should be 5 or higher.

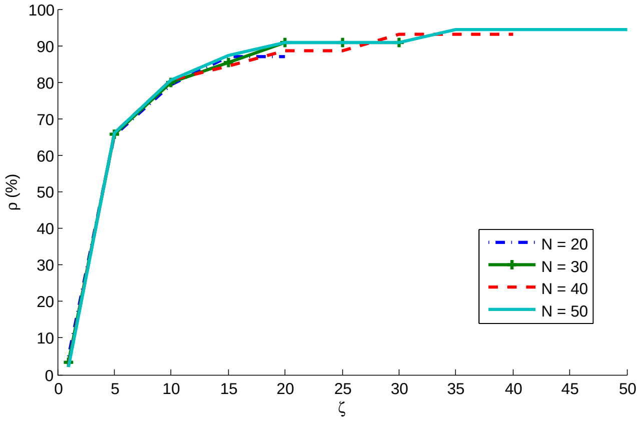


Figure 5.5: Sleep ratio of the nodes below the sink in WASP for a ζ-balanced tree.

5.3.2.4 Scalability

The number of nodes in a WBAN is limited by nature of the network. It is expected that the number of nodes will be in the range of 20–50, see Section 2.1. Our address structure, see section 5.3.1.5, supports up to 64 addresses which is most likely sufficient for WBANs. If more addresses need to be supported, the proposed address structure can be altered. Instead of 6 bits, 14 bits can be used. This will however negatively affect the amount of overhead generated by WASP.

Further, the more nodes in the network, the more data will be sent. This will negatively affect the maximum throughput per node.

5.3.2.5 Interference

Although WASP is slotted, interference can arise from nearby subtrees. The resulting interference can be minimized by randomizing WASP schemes. This randomization is not unlimited, e.g. the position of the contention slot is currently fixed, but it can reduce the interference probability.

5.3.3 Implementation and Validation

5.3.3.1 Implementation

The protocol was implemented in nsclick [34], a simulator that allows Click Modular Router [35] instances to run in the NS-2 network simulator.

Figure 5.6 gives an overview of the different Click elements and interconnections in our implementation. The packets coming from the network are dumped to NS-2 traces and then classified according to their type. Data is processed or forwarded, joins are handled during contention slots and the WASP-schemes are analyzed. If a node does not have a parent it will react on a scheme with a join, otherwise a new scheme is prepared.

Two extra elements, without any incoming or outgoing ports, are used: WASPTiming, responsible for the slot timing issues, and WASPInfobase which stores useful information needed by all elements in a node.

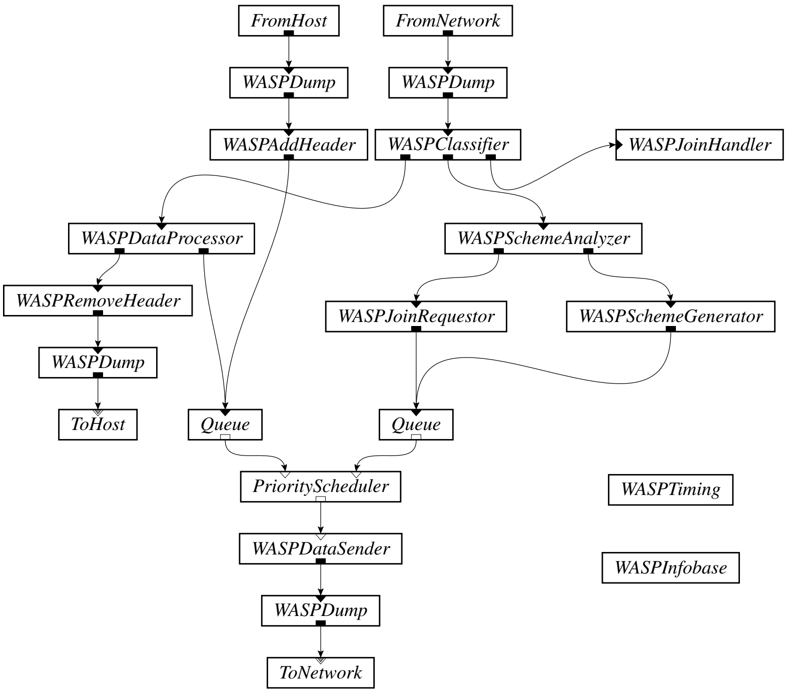


Figure 5.6: Overview of different click elements and interconnections

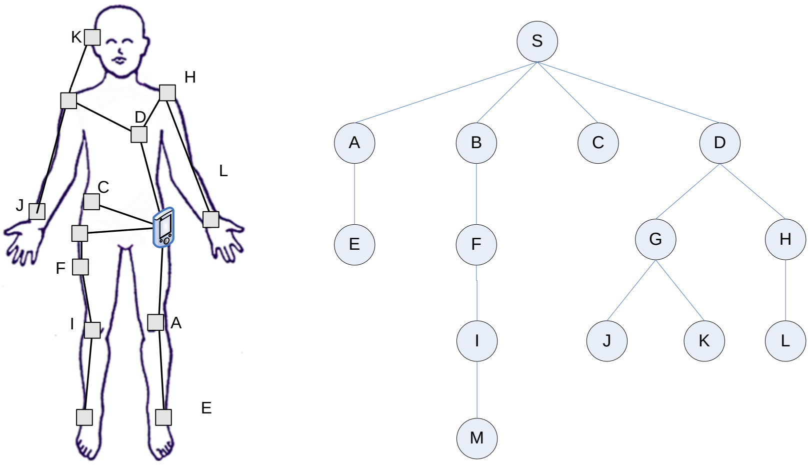
As mentioned in section 2.3.1.3, the path loss between two nodes on the body is highly different from the path loss in free space. Hence, we have adapted the propagation model of NS-2 to the path loss model of (2.4) which accurately models the path loss near a flat phantom for muscle tissue at 2.4 GHz. Further, we have used the energy model for the radio with the parameters of the Nordic transceiver, see Section 2.3.3.2. The bit rate is thus set to 1 Mbps.

5.3.3.2 Example Scenario

In this section, a more elaborate example is given and discussed. The network used is that of Figure 5.7, where 5.7(a) shows the network on a body and 5.7(b) a more generic view of the network in the form of a spanning tree. We assume that each node has data to send, which is reasonable as sensor traffic is normally constant bit rate traffic. The sequence of the nodes is not randomized in this example and all sensor data is sent to a sink. Each node sends a data packet of 500 bytes each 100 ms. As an example, (5.18) shows the WASP-scheme of the sink. The numbers correspond to the ones used in (5.1).

$$\overbrace{S}^{1}\ \underbrace{ABCD}_{2}\ \overbrace{.^{6}}^{3}\ \underbrace{ABBBDDDDD}_{4}\overbrace{X}^{5} - \underbrace{1111111111111}_{6} \tag{5.18}$$

From this scheme, it can be seen that the cycle length is 21 slots. This corresponds to the result of (5.8) when the parameters of this network is used. The maximum throughput calculated with (5.10) is 61,9% or 619 kbps.



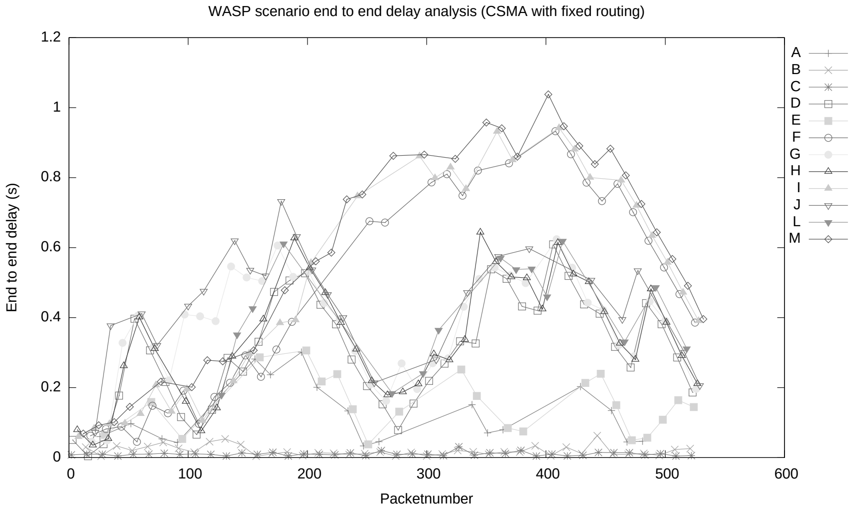
(a) Network on the body.

(b) Abstract view of the network.

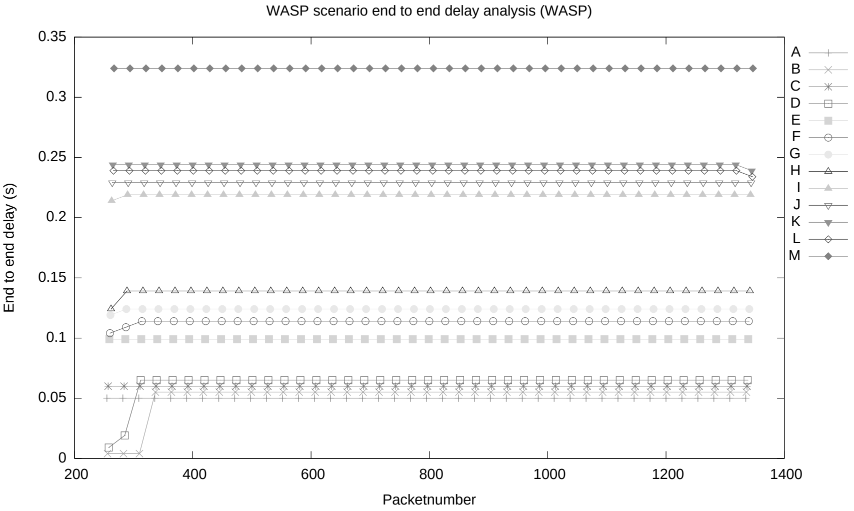
Figure 5.7: Simulated network.

Figure 5.8 shows the end-to-end delay for this reference scenario when running 7.5 seconds. Figure 5.8(a) shows the results when CSMA is used combined with fixed (optimal) routing. The delay shows high variation and regularly exceeds 0.35 seconds. About 30% of the packets are dropped and node K does not even succeed at transmitting a single packet. Figure 5.8(b) shows the results when WASP handles medium access and routing and the slotsize is 5 ms. The delay is fixed and the levels are clearly visible. The maximum delay is 0.324072 seconds and no packets are dropped. The smaller delays at the left of this graph can be explained by the absense of traffic forwarding in the beginning of the simulation. The larger number of packets is due to the WASP-schemes that are broadcast each cycle.

In this example, node M has the highest delay. This node is situated on level 4 and thus roughly needs 3 WASP-cycles, according to (5.14). Calculating the delay explicitly, we see that the delay only amounts to 43 time slots. This can be



(a) End-to-end delay with CSMA and fixed routing.



(b) End-to-end delay with WASP.

Figure 5.8: End-to-end delay comparison.

verified by writing down the whole WASP-cycle, which is omitted due to space constrictions. Node M sends its data to node I at time slot 12 in the first WASP-cycle. In the second cycle, node I forwards it to node F. And in the third cycle, node F forwards it to node B who forwards it to the sink in the same cycle. The packet arrives at the sink in time slot 14. Adding this up, we get 45 time slots or a delay of 225 ms. The extra delay of 100 ms comes down to the time difference between the generation of the packet and the availability of a slot for node M. More generally, this extra delay is a value between 0 and the length of the WASP-cycle, hence in this example at most 105 ms.

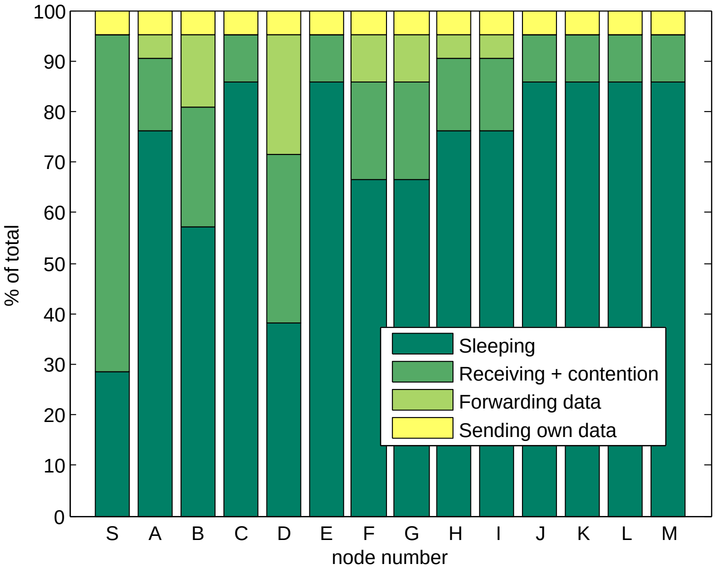


Figure 5.9: Time usage in a node. The sink has number 1, node A is number 2 and so on.

As said in section 5.3.2.2, the nodes can turn their radio off in slots where they are not involved in the communication. The time usage of the nodes is depicted in Figure 5.9. In the example scheme the sink can turn its radio off in the silent period as it will not receive any data from its child nodes. Thus, the radio can be turned of 6 slots. Node A can also turn its radio off in its silent period, when it knows that its siblings are sending and when none of its children are due to send data. So, node A can sleep 16 slots. The sleep ratios in this figure correspond to one calculated with (5.17).

5.3.4 Discussion

The WASP-protocol can be extended in order to even further ameliorate its performance. Different steps can be taken. For example, additional information can be advertised in the WASP-scheme, such as load, the quality of the link, etc. This can be used to choose a better parent and accordingly a better tree structure. Another method is data aggregation for lowering the number of transmissions. Doing this a node will put the data of several packets into one packet thus lowering the overhead per data bit sent.

The most important performance gain can be obtained by determining the optimal length of a time slot. This can be done based on the delay requirements for a specific application (5.14). A major improvement would be the extension of WASP to handle non-static networks. Currently almost no mobility is supported which is of course a requirement for a realistic wireless Body Area Network.

5.4 Conclusion

In this chapter we have given an overview of the existing MAC-strategies for WSNs and WBANs. In particular, the maximum throughput and minimum delay of IEEE 802.15.4 were determined. We have presented the exact formula for calculating the maximum throughput of the unbeaconed version of IEEE 802.15.4 for different frequency bands and scenarios. It was concluded that the throughput varies with the number of data bits in the packet. In the 2.4 GHz band a maximum throughput of 163 kbps or an efficiency of 64.9% can be achieved. The other frequency bands offer a higher efficiency, but a lower effective throughput. By changing the back off exponent, a higher throughput can be obtained. It is concluded that the bandwidth efficiency is rather low due to the small packet size imposed in the standard. Further, it was argumented that IEEE 802.15.4 is not the best solution for supporting communication in WBANs.

We have presented WASP, a new cross layer protocol for wireless Body Area Networks that both handles channel medium access and routing. For this purpose, a spanning tree is set up in a distributed manner and timeslots are used. Every node sends out a proprietary WASP-scheme to inform the nodes of the following level when they are allowed to send. These WASP-schemes are generated locally in each node. It is shown that the throughput efficiency can reach up to 94%, depending on the number of levels used and the size of the network. The end-to-end delay is shown to be fixed and related to the number of levels in the tree. WASP supports limited bidirectional traffic. Although WASP offers an answer to the challenges encountered in a WBAN, several improvements can be made in terms of energy efficiency and reliability, especially in supporting mobility. In order to counter these shortcomings, we propose CICADA in the following chapter.

References

[1] I. Demirkol, C. Ersoy, and F. Alagoz. *MAC protocols for wireless sensor networks: a survey*. IEEE Communications Magazine, 44(4):115–121, April 2006.

[2] P. Baronti, P. Pillai, V.. Chook, S. Chessa, A.Gotta, and Y. Fun Hu. *Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards*. Computer Communications, 30(7):1665–1695, May 2007.

[3] W. Ye, J. Heidemann, and D. Estrin. *An Energy-Efficient MAC protocol for Wireless Sensor Networks*. In Proceedings of the IEEE Infocom, pages 1567–1576, New York, NY, USA, June 2002. USC/Information Sciences Institute, IEEE.

[4] T. van Dam and K. Langendoen. *An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks*. In SenSys03, pages 171–180, Los Angeles, CA, November 2003.

[5] A. El-Hoiydi and J.-D. Decotignie. *WiseMAC: an ultra low power MAC protocol for the downlink of infrastructure wireless sensor networks*. In ISCC '04: Proceedings of the Ninth International Symposium on Computers and Communications 2004 Volume 2 (ISCC"04), pages 244–251, Washington, DC, USA, 2004. IEEE Computer Society.

[6] J. Polastre, J. Hill, and D. Culler. *Versatile low power media access for wireless sensor networks*. pages 95–107, Baltimore, MD, November 2004.

[7] M. Buettner, G. Yee, E. Anderson, and R. Han. *X-MAC: A Short Preamble MAC Protocol For Duty-Cycled Wireless Networks*. Technical report cu-cs-1008-06, University of Colorado at Boulder, May 2006.

[8] W. Ye, F. Silva, and J. Heidemann. *Ultra-low duty cycle MAC with scheduled channel polling*. In SenSys '06: Proceedings of the 4th international conference on Embedded networked sensor systems, pages 321–334, New York, NY, USA, 2006. ACM Press.

[9] G. Lu, B. Krishnamachari, and C. S. Raghavendra. *An adaptive energy-efficient and low-latency MAC for data gathering in wireless sensor networks*. In Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International, April 2004.

[10] I. E. Lamprinos, A. Prentza, E. Sakka, and D. Koutsouris. *Energy-efficient MAC Protocol for Patient Personal Area Networks*. In Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the, pages 3799–3802, Shanghai,, 2005.

[11] E. Farella, A. Pieracci, L. Benini, and A. Acquaviva. *A Wireless Body Area Sensor Network for Posture Detection*. In ISCC '06: Proceedings of the 11th IEEE Symposium on Computers and Communications, pages 454–459, Washington, DC, USA, 2006. IEEE Computer Society.

[12] O. C. Omeni, O. Eljamaly, and A. J. Burdett. *Energy Efficient Medium Access Protocol for Wireless Medical Body Area Sensor Networks*. In Medical Devices and Biosensors, 2007. ISSS-MDBS 2007. 4th IEEE/EMBS International Summer School and Symposium on, pages 29–32, Cambridge, UK,, August 2007.

[13] Hu. Li and J. Tan. *Heartbeat driven medium access control for body sensor networks*. In HealthNet '07: Proceedings of the 1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments, pages 25–30, New York, NY, USA, 2007. ACM.

[14] IEEE 802.15.4-2003: IEEE Standard for Information Technology - Part 15.4: Wireless Medium Access Control and Physical Layer specifications for Low Rate Wireless Personal Area Networks.

[15] ZigBee Alliance, official webpage: http://www.zigbee.org.

[16] P. Johansson, M. Kazantzidis, R. Kapoor, and M. Gerla. *Bluetooth: an enabler for personal area networking*. IEEE Network, 15(5):28–37, September/October 2001.

[17] B. Latré, P. De Mil, I. Moerman, N. Van Dierdonck, B. Dhoedt, and P. Demeester. *Maximum Throughput and Minimum Delay in IEEE 802.15.4*. Lecture Notes in Computer Science , Proceedings of the 1st International Conference on Mobile Ad-hoc and Sensor Networks, 3794:866–876, December 2005. ISSN 1796-2056.

[18] B. Latré, P. De Mil, I. Moerman, N. Van Dierdonck, B. Dhoedt, and P. Demeester. *Throughput and Delay Analysis of Unslotted IEEE 802.15.4*. Journal of Networks, 1(1):20–28, May 2006. ISSN 1796-2056.

[19] N. F. Timmons and W. G. Scanlon. *Analysis of the performance of IEEE 802.15.4 for medical sensor body area networking*. In Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on, pages 16–24, October 2004.

[20] H. Li and J. Tan. *An Ultra-low-power Medium Access Control Protocol for Body Sensor Network*. In Engineering in Medicine and Biology Society,

2005. IEEE-EMBS 2005. 27th Annual International Conference of the, pages 2451–2454, Shanghai,, 2005.

[21] N. Golmie, D. Cypher, and O. Rebala. *Performance analysis of low rate wireless technologies for medical applications*. Computer Communications, 28(10):1266–1275, June 2005.

[22] D. Cavalcanti, R. Schmitt, and A. Soomro. *Performance Analysis of 802.15.4 and 802.11e for Body Sensor Network Applications*. In 4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007), volume Volume 13, pages 9–14. Springer Berlin Heidelberg, 2007.

[23] IEEE 802.15 WPAN Task Group 6 Body Area Networks.

[24] K. Akkaya and M. Younis. *A survey on routing protocols for wireless sensor networks*. Ad Hoc Networks, 3(3):325–349, 2005.

[25] Hongliang Ren and Max Q. H. Meng. *Rate Control to Reduce Bioeffects in Wireless Biomedical Sensor Networks*. In Mobile and Ubiquitous Systems - Workshops, 2006. 3rd Annual International Conference on, pages 1–7, San Jose, CA,, July 2006.

[26] Q Tang, N. Tummala, S. K. S. Gupta, and L. Schwiebert. *Communication scheduling to minimize thermal effects of implanted biosensor networks in homogeneous tissue*. IEEE Transactions on Biomedical Engineering, 52(7):1285–1294, July 2005.

[27] A. Bag and M. A. Bassiouni. *Energy Efficient Thermal Aware Routing Algorithms for Embedded Biomedical Sensor Networks*. In Mobile Adhoc and Sensor Systems (MASS), 2006 IEEE International Conference on, pages 604–609, Vancouver, BC,, October 2006.

[28] D. Takahashi, Y. Xiao, and F. Hu. *LTRT: Least Total-Route Temperature Routing for Embedded Biomedical Sensor Networks*. In IEEE Globecom 2007, November 2007.

[29] T. Watteyne, S. Augé-Blum, M. Dohler, and D. Barthel. *AnyBody: a Self-organization Protocol for Body Area Networks*. In Second International Conference on Body Area Networks (BodyNets), Florence, Italy, 11-13 June 2007. 2007.

[30] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. *Energy-efficient communication protocol for wireless microsensor networks*. In System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on, January 2000.

[31] M. Moh, B. J. Culpepper, Lan Dung, Teng-Sheng Moh, T. Hamada, and Ching-Fong Su. *On data gathering protocols for in-body biomedical sensor networks*. In Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE, volume 5, November/December 2005.

[32] A. G. Ruzzelli, R. Jurdak, G. M.P O'Hare, and P. Van Der Stok. *Energy-efficient multi-hop medical sensor networking*. In HealthNet '07: Proceedings of the 1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments, pages 37–42, New York, NY, USA, 2007. ACM.

[33] Nordic, nRF 2401 datasheet [online] http://www.nordicsemi.com /index.cfm?obj=product&act=display&pro=64.

[34] M. Neufeld, A. Jain, and D. Grunwald. *Nsclick:: bridging network simulation and deployment*. In MSWiM '02: Proceedings of the 5th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems, pages 74–81, New York, NY, USA, 2002. ACM.

[35] E. Kohler, R. Morris, B. Chen, J.Jannotti, and M. F. Kaashoek. *The click modular router*. ACM Trans. Comput. Syst., 18(3):263–297, 2000.

6

CICADA: Energy Efficient and Reliable Networking

In this chapter, we present CICADA (*Cascading Information retrieval by Controlling Access with Distributed slot Assignment*). This is a low energy protocol designed for wireless, multi-hop, mobile Body Area Networks and can be considered as a large improvement of WASP that was discussed in Chapter 5. Data gathering trees are autonomously set up and used to route data from the nodes towards the personal device or sink. In each cycle, the tree structure is used to allocate time slots to the different nodes in a distributed manner. CICADA breaks a cycle down in a control subcycle and a data subcycle, allowing lower delay, lower packet loss and better robustness against mobility.

6.1 Protocol Description

6.1.1 General Principles

CICADA can be seen as a cross-layer protocol as it uses the same packets to take care of medium access as well as routing. These packets are also used to detect the presence or absence of the children and to control medium access. The protocol sets up a spanning tree and divides the time axis in slots in order to lower the interference and avoid idle listening. The assignment of the slots is done in a distributed manner and slot synchronization is possible as a node knows the length of each cycle. Each node informs its children (i.e. the nodes just beneath it) when they are allowed to send their data. Routing itself is not complicated in CICADA anyway as data packets are routed up the tree.

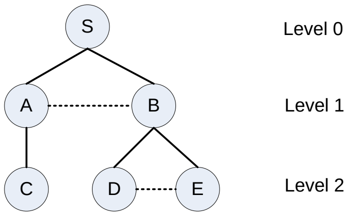


Figure 6.1: An example of the network topology. The full lines indicate the tree structure. The communication channels are indicated with both the full and dotted lines.

Data transfer is defined by a sequence of cycles. At the beginning of each cycle, the slots in the remainder of the cycle are assigned. Each parent sends a SCHEME-message to all their children, containing their slot allocation scheme. A node calculates the scheme of its children based on the scheme it has received from its parent. Each cycle is divided in two parts: the control subcycle and the data subcycle. Each subcycle has its own slot allocation scheme: the control scheme and the data scheme respectively. These schemes are both sent in the control subcycle. When all nodes have received their scheme, the control cycle has ended and the data cycle starts.

Each node is assigned one slot in the control subcycle. As slots in the control subcycle are only used to send the short SCHEME-messages, slots can be shorter, e.g. by a factor 5 or more. When a node has received such a SCHEME-message in a control slot, it can turn its radio off as no more packets will arrive in the control subcycle. The data subcycle is used to forward the data from the nodes to the sink. Unlike in WASP, the first nodes to start sending data are the nodes at the bottom of the tree. Doing so, all data can be sent to the sink in one cycle. This lowers the end-to-end delay tremendously.

Thus, as can be seen in Figure 6.2, control information is sent downwards from the sink to all nodes in the control subcycle. In the data subcycle, all data is sent upwards to the sink. In the following, we will discuss both cycles in detail using the example network from Figure 6.1. The tree is set up in such a way that communication is only possible between a child and its parent or between siblings.

6.1.2 Control Subcycle

The control subcycle is used for transferring the schemes (i.e. the control scheme and the data scheme) to all nodes. At the start of the control subcycle, the sink

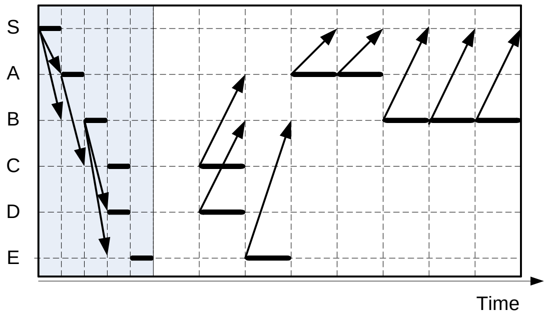


Figure 6.2: Communication in the example network of Figure 6.1. The arrows indicate the transmission direction. The black boxes show when the node is transmitting. The shaded area is the control cycle.

sends the first message (i.e. its SCHEME-message). Table 6.1 shows which nodes are allowed to send in which slot for the example network. The assignment of the slots in the control subcycle is done using the control scheme. Each control scheme contains the following information of the control subcycle:

Slot	**1**	**2**	**3**	**4**	**5**
	S	A	B	C + D	E

Table 6.1: Steady-state control subcycle for the example network of Fig. 6.1

- The control scheme indicates the order in which the children are allowed to send their control scheme;
- The total length of the control subcycle in timeslots T_{CC}, starting from the transmission of the control scheme of the sink. Stated otherwise, this is the total number of slots needed to allow all devices to send their scheme. In the example, the length is 5;
- The remaining slots of the control subcycle, including the slot in which the node is transmitting. This is needed to know when the data subcycle begins.

This information is used to calculate the exact slot at which the node may send. For example, node S sends its scheme with the following information:

- *control scheme* = AB
- *control subcycle length* T_{CC} = 5
- *remaining slots* = 5

Nodes A and B receive this information. Node A sees in the control scheme that it is allowed to send first, so it will send its SCHEME-message in the following slot containing the control scheme, the additional information and the data scheme. Node B will send in the slot thereafter. The control subcycle length is the same. However, as node C cannot send simultaneously with node B, node A will add a wait slot to its control scheme which becomes "$.C$". Thus, in the control cycle, node C has a waiting period of length 1. Node B will send the following: control scheme DE, control subcycle length 5 and 3 remaining slots. The remaining length indicates how many slots are left in the control subcycle and is thus used to know the start of the data subcycle. Table 6.2 gives an overview of the control subcycle information of the different nodes.

	S	**A**	**B**	**C**	**D**	**E**
remaining length	5	4	3	2	2	1
control scheme	AB	$.C$	DE	NA	NA	NA

Table 6.2: Control subcycle information of the nodes. NA indicates that no control scheme is available (i.e. the node has no children).

6.1.3 Data Subcycle

For the example network in Figure 6.1, the data schemes of the nodes, or stated otherwise the division of the time slots in the data subcycle, can be seen in Table 6.3.

Slots	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
S	.	.	.	.	A	A	B	B	B	X
A	.	C	X							
B	.	D	E	X						
C	X									
D	X									
E	X									

Table 6.3: Steady-state data subcycle for the example network. The dots represent waiting slots. The overview only shows when the radio is receiving data from a child node and the contention slot.

The data scheme consists of three parts: a receiving period (length α), a waiting period (length β) represented by dots and a sending period. In the waiting period, the node must remain silent and should turn off its radio. In the receiving period, the node receives data from its children and in the sending period the node sends data to its parent. For example, node B has a waiting period of 1 slot, a receiving period of 2 slots and a sending period of 3 slots. The data scheme determines when the child nodes are allowed to transmit in the receiving period. The last slot of each data scheme is a contention slot which is used to allow new children to join the network, see section 6.1.4.

Each node maintains a table for its children containing the following information:

α_i The number of data slots needed to receive the data from node i

β_i The number of data slots node i needs to receive data from its children, the length of the waiting period of node i and 1 slot for contention.

This table is called the ChildTable. Each child is granted the number of data slots indicated in the ChildTable (α_i). Each time a node sends a data packet, a small amount of additional information is put in the data header. This header contains α_n and β_n, assuming that node n is sending. These values are calculated as follows.

$$\alpha_n = \sum_{i \in Ch_n} \alpha_i + \delta_n \tag{6.1}$$

$$\beta_n = \max_{i \in Ch_n} \beta_i + \sum_{i \in Ch_n} \alpha_i + 1 \tag{6.2}$$

In these formulas, Ch_n represents the children of node n and δ_n represents the number of data slots that is needed for node n to transmit its own generated data. The $\max$ function is needed as different branches of the tree are allowed to send simultaneously. When a parent receives a packet from its child, it will extract this information from the header and update its ChildTable. Each child has to send this additional information each data subcycle. If a child has no data packet to send or to forward, it will send a HELLO-message to its parent containing only that information. Doing so, the parent knows that the child is still connected to the tree. In table 6.4, the ChildTables for nodes S, A and B are given. As nodes C, D and E have no children, their ChildTables will be empty.

S

	α_i	β_i
A	2	3
B	3	4

A

	α_i	β_i
C	1	1

B

	α_i	β_i
D	1	1
E	1	1

Table 6.4: ChildTables

Based on the information in the ChildTable, the node can calculate the data scheme as follows. First, it determines how long the child nodes have to wait until they can start sending by taking the maximum of the children's waiting period:

$\max_{i \in Ch_n} \beta_i = \beta$. Doing so, the structure of the tree and the subsequent simultaneous transmission are taken into account. Then each child node i is granted α_i slots, leading to a receiving period of $\alpha = \sum_{i \in Ch_n} \alpha_i$ slots. At the end, a contention slot is added.

6.1.4 Joining the Network

In each data subcycle, a contention slot is included to allow nodes to join the tree. A new node listens for SCHEME-messages. When one is received, a JOIN-REQUEST message is sent in the contention slot, preferably after a random delay within the contention slot to avoid contention of simultaneously sent JOIN-REQUEST messages. This JOIN-REQUEST message also contains the number of slots a node needs to transmit its own generated data (δ_i), to enable immediate optimal resource allocation. When the parent hears the join message, it will include the node in the next cycle by updating its ChildTable. As it is assumed that a node i joining the network has no children, δ_i equals α_i and β_i equals 1.

6.1.5 Detection of Connection Loss

Each node will send at least two packets per cycle: a data packet or a HELLO-message and a control packet[1]. If a parent does not receive a packet from a child for two or more consecutive cycles, the parent assumes that the child is lost and removes it from its ChildTable. If a child does not receive packets from its parent for two or more consecutive cycles, the child will assume that the parent is gone and will try to join another node.

6.1.6 Supporting Bidirectional Traffic

In a WBAN the traffic from the personal device mainly consists of control data such as settings for the devices. In some applications, two nodes will need to communicate to each other, such as a sensor node that measures the glucose level in the blood and an actuator that injects insulin. The communication between these devices will be limited to a few byte per second. As already mentioned, the two subcycles in CICADA define two directions of traffic: the control cycle from the personal device towards the sensor nodes and the data cycle from the sensor nodes to the personal device. In the control cycle, additional data can be added to the schemes and propagated from the sink downwards the tree. Doing so, (asymmetric) bidirectional traffic can be supported. The communication between two devices can occur in two ways: via the sink or not. In the latter, the device sends its data to the sink, the sink notices that the information is destined for another device and sends it to the device in the control cycle. In the former, the

[1] The SCHEME-message are also referred to as control packets.

message is propagated in the tree until a common ’grandparent’ is found. This would require adding a lightweight routing table to the devices, i.e. a table with the ID’s of the devices below it.

6.1.7 Traffic Awareness

CICADA is built in such a way that it supports heterogeneity of the nodes. Nodes can ask the exact number of slots at any time by changing the value of δ_n in (6.1). In the following cycle, these extra slots will be allocated to the node. If desired, the node can remain connected and forward data from other nodes. Doing so, the node cooperates in the network and spreads the load of the network. If the node has data to send, an additional slot is requested.

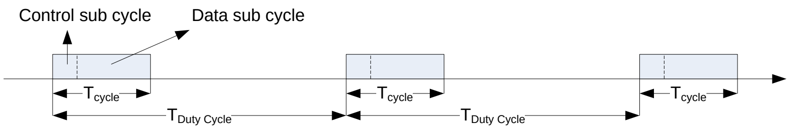


Figure 6.3: The succession of an active period and a sleeping period in CICADA. The ratio of T_{cycle} and $T_{DutyCycle}$ is called the duty cycle (Δ).

CICADA has been developed to support high-traffic Body Area Networks where delays should be low, i.e. all sensors send data often instead of buffering it locally. When used for low-traffic networks, an optimization is possible. The lifetime of the network can be further improved by using an *Adaptive Duty Cycle*. When the load of the network is low, it is not necessary to schedule the cycles immediately after each other. Instead, a sleeping period can be added between two cycles. The succession of an active period (containing one cycle) and a sleeping period is called a duty cycle period. An example is given in Figure 6.3. The length of one cycle is T_{cycle} and the length of the duty cycle period is $T_{DutyCycle}$. The ratio of T_{cycle} and $T_{DutyCycle}$ is called the duty cycle Δ. The lower Δ, the more energy efficient the system will be (as the devices sleep longer), but the lower the throughput. According to the load of the network, the duty cycle can be adapted on the fly. This is determined by the sink who can make a decision based on the number of slots reserved by each node.

6.1.8 Using a Data Gathering Tree

The shape of the tree has a strong influence on the efficacy of the protocol, as will be pointed out in Section 6.2. In the current implementation, the tree is formed depending on link formation times, i.e. the node heard first becomes the parent. A possible adaptation might consist of adding extra information to the SCHEME-message. This can include parameters like signal strength, load, remaining energy, etc.

In literature, several tree forming algorithms are found for WSNs. In PEGASIS [1], the nodes are organized to form a chain. The nodes take it in turn to transmit data to the base station. In every round a node is elected that collects the data from other nodes along the chain and transmits it to the sink. To allow for efficient use of this technique, a threshold is put on the distance of the elected node from the base station. Doing so, the energy required to send the data to the sink is averaged out over all the nodes. D-MAC [2] sets up a data gathering tree by letting each node choose from its neighbors the node closest to the sink as the next hop. In [3] a breadth-first search (BFS) algorithm is used to form a convergecast tree. They further prove that a spanning tree constructed via the BFS-algorithm has no conflicting edges, i.e. edges that send simultaneously. In [4] a combination of TDMA and CDMA is used to limit the interference between the branches. This is however too complex to be used in a WBAN-device.

Supporting mobility in CICADA can be defined as transforming one tree into another one. Due to the fast detection of movements (see Section 6.1.5) mobility becomes feasible.

Nodes higher in the tree (i.e. near the personal device) relay a lot for other nodes, requiring a lot of energy. This is caused by the tree-based network structure. However, if a node really cannot relay data because of energy shortage, it should not act upon joining nodes or drop children. Those nodes will then look for another parent and spare the energy-constrained node. Another possible solution is the use of dedicated relay nodes near the top of the tree, i.e. near the sink, as was done in Section 3.4.

6.1.9 Data Aggregation

Due to the tree structure of CICADA, a device receives messages from several nodes below it in the tree that need to be forwarded. Some of this sensor data is only a few bytes long, e.g. temperature, blood pressure and so on. These messages will only take up a small portion of the available time slot. An extension to CICADA is to aggregate the messages and transmit them as a single packet. Thus, a device that has more than one message to send or forward, aggregates and sends the messages in one packet during the same time slot, creating a data gathering tree. Doing so, the transmission overhead is reduced and energy is saved. The number of messages that can be aggregated in one packet varies according to the maximum length of the packet which in his turn depends on the length of the time slot. Messages are only aggregated as long as the packet size does not exceed the maximum packet length. Aggregating messages is beneficial for the lifetime of the network, as was shown in Chapter 3.

This mechanism will however change the way how a node calculates its waiting and receiving period. Indeed, a node n can no longer calculate its waiting period with (6.1) as the data from its children will be aggregated. In order to know the length of its waiting period, it has to know the length of the packets from its children. Therefore the following changes will be applied to the protocol when aggregation is used.

When a node n receives a packet from node i, it extracts its α_i and β_i and it extracts the length of the last data packet received from node i. This length in bits is referred to as κ_i. For example, say node i has α_i = 3, thus 3 data packets will be sent to node n. The first 2 packets will be full, so the data from the third packet can be used for aggregation at the parent node n. The value of κ_i will also be stored at the ChildTable of node n. The following formula can now be used for calculating the waiting period:

$$\alpha_n^A = \sum_{i \in Ch_n} (\alpha_i - 1) + \left\lceil \frac{\kappa_n + \sum_{i \in Ch_n} \kappa_i}{\kappa_{max}} \right\rceil \tag{6.3}$$

where κ_n is the length of the own data and κ_{max} the maximum packet length.

The calculation of the receiving period β_n^A remains the same as (6.2). Each node will still need α_i slots for sending its data, the maximum does not change and 1 contention slot is added.

By using data aggregation, the cycle length T_{CC} will be lower. The performance of CICADA will improve as the nodes will need lesser time to send and forward the data. Further, when the data rates are heterogeneous, the slot length can be bigger without causing extra delays in the network. The slots are also more efficiently used as more data packets are sent in one slot.

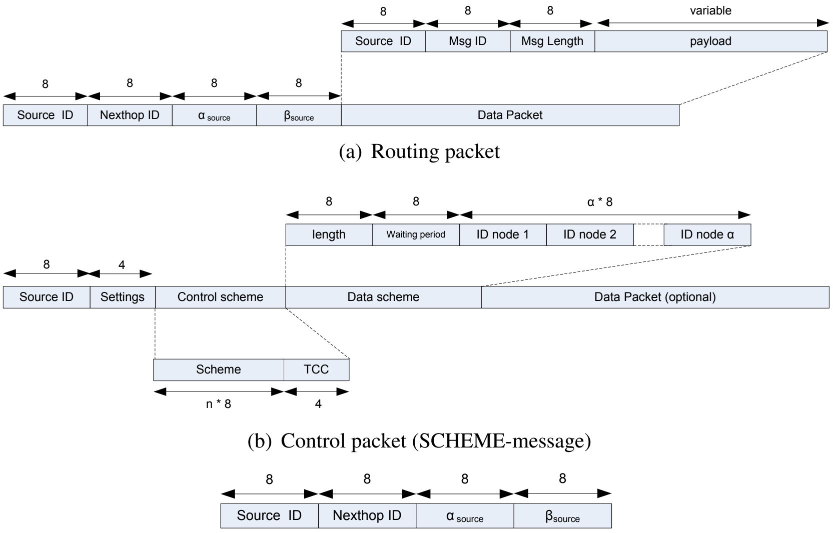
6.1.10 Using Relay Devices

Relay devices can be easily added by letting the device join the network and setting the value of δ_n to zero. They can be placed anywhere in the network, but as was concluded in Chapter 3 they are best placed not far from the personal device. Further, the relay devices can publish their presence in the SCHEME-message. Doing so, the normal nodes can choose a relay device as parent instead of a normal node.

6.1.11 Packet Format

The packet format in CICADA depends on the type of message sent. In general, three types can be distinguished: a control packet, a routing packet and a JOIN-REQUEST-message. The length of the node ID's is limited to 8 bits, allowing a network of 255 nodes. This is sufficient for a WBAN.

Figure 6.4 shows the packet format for the different messages. The routing packet is used for sending data to the next hop in the data subcycle. It contains the



(a) Routing packet

(b) Control packet (SCHEME-message)

(c) JOIN-REQUEST-message

Figure 6.4: Packet format in CICADA. The numbers indicate the length in bits. α_{source} is the α_n of node n sending the packet.

ID of the sending node (say node n) and the ID of the parent supposed to receive the message. Further, the node sends its α_n and β_n to the parent, calculated by (6.1) and (6.2) respectively, followed by a data packet. This packet contains the ID where the data was originated, a message ID and the payload of the data. The length of the payload is variable and limited by the maximum packet size. The control packet contains the ID of the node sending it, followed by the control scheme and the data scheme. If bidirectional traffic is supported and data is sent during the control cycle, the `settings`-bit is set to 1. The data packet is added after the data scheme. The JOIN-REQUEST message only contains the ID of the sending node, the ID of the desired parent (Nexthop ID) and its α_n and β_n. The HELLO-message is similar to the JOIN-REQUEST message.

From this packet structure, it can be seen that the length of the control packet is limited to a few bytes, depending on the number of slots the children need to send their data ($\alpha = \sum_{i \in Ch_n} \alpha_i$). The length of a slot in the control cycle can thus be smaller than a data slot. If the node with the most data traffic has α slots for receiving the data and n children, the length of the control packet is $(4+n+\alpha)$ bytes. If we further assume that the maximum length of a data packet is 512 bytes, then the control slot length can be more then ten times smaller if α and n are not higher than 20.

6.1.12 Differences with WASP

In CICADA all traffic reaches the sink in 1 cycle, which results in a lower delay. Each CICADA-cycle consists of 2 different subcycles: the control subcycle and the data subcycle. This results in enhanced mobility support: it is possible to leave a parent, to detect loss of a parent or a child and to join a node with much lower delays. CICADA does not rely on broadcasting to pass information from children to their parents, instead additional information is put in the data packets or HELLO packets are used. Moreover, generating CICADA schemes is easier because of the simple computation of the waiting period and the receive period. This is important as CICADA is meant to run on sensors where computational resources are scarce. The performance differences will be shown in the next session.

6.2 Performance Analysis

In this section, we present the performance analysis of CICADA. We will address performance issues such as the energy efficiency, reliability and mobility. Further, it is assumed that the nodes are loosely synchronized. The SCHEME-messages can be extended with a time stamp that can be used to adjust the clock of the child nodes.

6.2.1 Energy Efficiency

The most important causes of wasting energy in radio communication are idle listening, overhearing and collisions. CICADA takes care of all those causes by assigning slots in the control cycle and using them in the data cycle. Nodes only need to operate their radio during used slots, i.e. the slots where the node is receiving or sending data and during the control subcycle. All slots are allocated so a node perfectly knows when it is allowed to sleep, when it has to send or when it has to switch on his radio to receive data. Idle listening and overhearing can occur in the control subcycle as the nodes have to wait for the control scheme of their parent and consequently have to switch on their radios. However, the slots of the control cycle are shorter than the ones of the data cycle and nodes can sleep when a scheme from their parent has been received. In the data subcycle, nodes only have to wake up when transmitting or receiving data. Using these mechanisms, the dissipation of energy is minimized.

A diagram of the different states and their transitions can be found in Figure 6.5.

The sleep ratio ρ_i of node i for a full duty cycle can be written as:

$$\rho_i = \frac{T_{cycle} - T_{on,i}}{T_{cycle}} \tag{6.4}$$

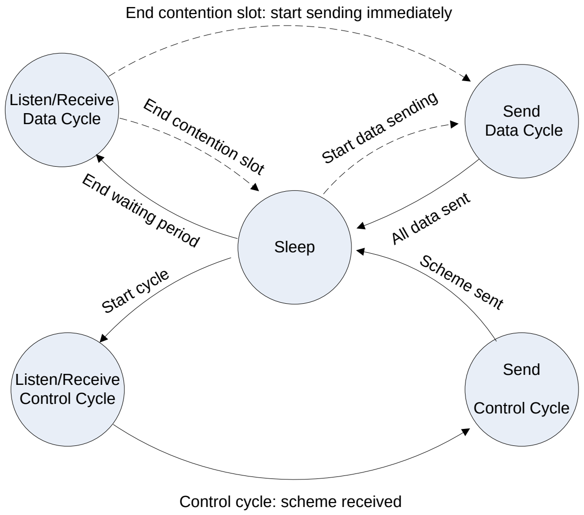

End contention slot: start sending immediately
Listen/Receive Data Cycle
End contention slot
Start data sending
Send Data Cycle
End waiting period
Sleep
All data sent
Scheme sent
Start cycle
Listen/Receive Control Cycle
Send Control Cycle
Control cycle: scheme received

Figure 6.5: State diagram in CICADA. The striped line pattern indicates two alternative paths, depending on the fact if the node directly starts sending after its contention slot or not. The radio is powered down in the sleep state.

$T_{on,i}$ is the time in slots a node i has its radio on. This corresponds to the exterior states in Figure 6.5.

$$T_{on,i} \;=\; T_{cc} \;+\; 2 \cdot \sum_{j \in Ch_i} \alpha_j \;+\; \delta_i \;+\; 1 \tag{6.5}$$

In this equation, T_{CC} represents the length of the control subcycle, the factor 2 in the second term indicates the time spent for receiving and sending the data, the third term covers the transmission of its own data and the fourth term the contention slot.

The cycle length T_{cycle} can be calculated using the length of the control subcycle, the waiting period and receiving period of the sink and a contention slot at the end:

$$T_{cycle} \;=\; T_{CC} \;+\; \sum_{i \in Ch_S} \alpha_i \;+\; \max_{i \in Ch_S} \beta_i \;+\; 1 \tag{6.6}$$

The last term can be rewritten as follows: let $\max_{i \in Ch_S} \beta_i = \gamma(S)$. The function $\gamma(n)$ gives the number of total waiting slots for a node n. In a data subcycle, the lowest level only has a contention slot and no waiting period, thus in the last but one level, $\gamma(n)$ should be 1.

$$\gamma(n) \;=\; \begin{cases} 1, & \bigcup\limits_{i \in Ch_n} V_i = \phi \\ \max\limits_{\forall i \in Ch_n} \big(\sum\limits_{j \in Ch_i} \alpha_j + \gamma(i) \;+\; 1 \big), & \bigcup\limits_{i \in Ch_n} V_i \neq \phi \end{cases} \tag{6.7}$$

The network topology clearly plays an important role in the energy efficiency of the protocol: it determines the cycle length as it effects the waiting period and the length of the control subcycle. If the tree structure is rather flat, the waiting period will be short and vice versa for a line topology. As in Section 5.3.2, we assume a ζ-balanced tree and that every node just needs one time slot per cycle (δ_i = 1 for all nodes). Doing so $\gamma(n)$ is given as

$$\gamma(n) = \sum_{i=0}^{L-2-y} \sum_{j=0}^{i} \zeta^j \tag{6.8}$$

where L is the number of levels in the network and y the level node n resides in (the sink has level 0). This can be proved as follows. The summation in (6.7) can be seen as the number of nodes below node i as every packet needs only one time slot. In a ζ-balanced tree, we thus can write $\sum_{j=1}^{l} \zeta^j$ where l depends on the level of node i. Together with the contention slot, which equals one time slot, we get the second summation of (6.8). Each level, except the last one, adds to the waiting period. Thus the first summation counts till $L-2-y$ as the first level has number 0.

In a ζ balanced tree, the length of the control subcycle T_{CC} can be written as $(L-1)\cdot\zeta+1$. The sink uses one slot and in each level below the sink all the children send their schemes one after the other before the next level starts sending. Of course, this is only the case when the tree is fully populated, i.e. when all the nodes on the last but one level have ζ children. In order to account for this overestimation, the number of nodes in the lowest level can be calculated with $N - \sum_{i=0}^{L-2} \zeta^i$. Further, we assume that all the nodes are equally divided in the lowest level: the number of nodes is thus divided by ζ and rounded to the nearest greater integer. The slots in the control subcycle can be smaller than the slots in the data subcycle, the ratio of the slot length in the data subcycle and control subcycle is called ϵ. Hence, T_{CC} becomes

$$T_{CC} = \frac{(L-2)\cdot\zeta + 1 + \left\lceil \frac{N - \sum_{i=0}^{L-2} \zeta^i}{\zeta} \right\rceil}{\epsilon} \tag{6.9}$$

The number of slots needed for receiving and forwarding the data for a node on the level below the sink is given by

$$\sum_{j \in Ch_i} \alpha_j = \left\lceil \frac{(N-1)-\zeta}{\zeta} \right\rceil \tag{6.10}$$

The only thing that we need now is the number of levels in the network. When we have a fully ζ-balanced tree where also the nodes of the last but one level have ζ children, we can write $N = \sum_{i=0}^{L-1} \zeta^i$. This can be rewritten as

$$L = \left\lceil \frac{\ln(N\cdot(\zeta-1)+1)}{\ln\zeta} \right\rceil \tag{6.11}$$

where the ceiling function is needed when the tree is not fully balanced, i.e. when there are not enough nodes to completely fill the lowest level. Now the sleep ratio for a node on the level below the sink for a ζ balanced tree can be formulated as

$$\rho = \frac{N + \sum_{i=0}^{L-2} \sum_{j=0}^{i} \zeta^j - 2 \cdot \left\lceil \frac{(N-1)-\zeta}{\zeta} \right\rceil - 2}{\frac{(L-1)\cdot\zeta+1+\left\lceil \frac{N-\sum_{i=0}^{L-2} \zeta^i}{\zeta} \right\rceil}{\epsilon} + N + \sum_{i=0}^{L-2} \sum_{j=0}^{i} \zeta^j + 1} \tag{6.12}$$

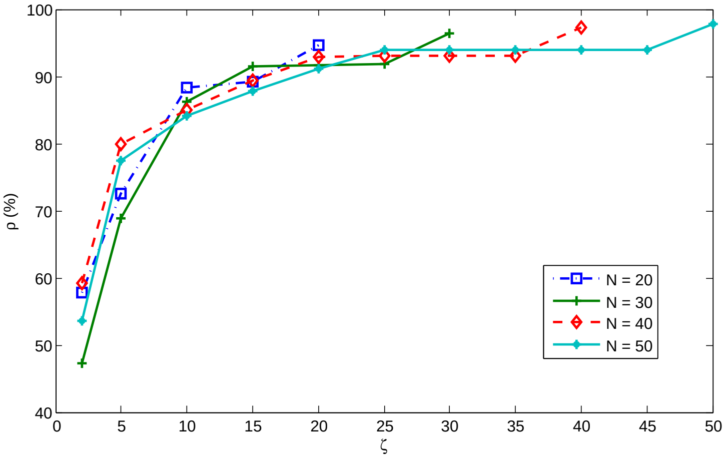


Figure 6.6: Sleep ratio in CICADA for varying ζ, varying network size N and full duty cycle. The slots in the control cycle are a factor 5 smaller.

Figure 6.6 compares the sleep ratio ρ for varying ζ for the nodes in level 1, below the sink. The sleep ratio is lower when ζ is smaller, similar to the sleep ratio in WASP. Further, the sleep ratio is almost similar for the varying network sizes. When ζ is higher than $N/2$, the sleep ratio remains stable as both the numerator and denominator change at the same rate. Every node forwards at most data from 1 child node and the length of the control cycle remains the same. For a network of 50 nodes, a sleep ratio of almost 93% is obtained. When ζ = N, a maximum is reached as all the nodes are directly connected to the personal device. Figure 6.7 shows the influence of shortening the timeslots in the control cycle. If the timeslots are five times shorter, the sleep ratio drops about 5 to 10%, especially for larger ζ. So, it is good practice to use a slot ratio of 10 or more in order to have an adequate energy efficiency. The figure also compares the sleep ratio of CICADA with WASP's. When ζ is small, CICADA outperforms WASP and for larger ζ, WASP performs better when the time slot ratio is low.

Overall, the tree should not have too many levels. That way, the nodes can sleep more and the upper bound for the delay will be lower, see [5] for more information.

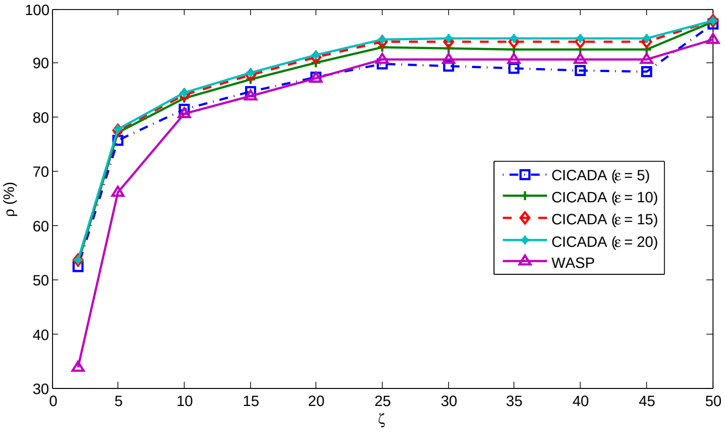


Figure 6.7: Sleep ratio in CICADA and WASP for varying ζ, N = 50 and full duty cycle. For CICADA, the ratio between the slot length in the data subcycle and control subcycle is varied (ϵ).

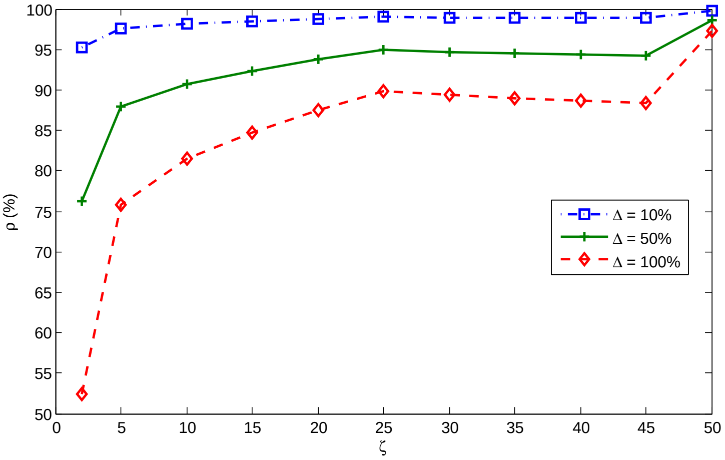


Figure 6.8: Influence of duty cycle on the sleep ratio in CICADA for varying ζ and N = 50.

The analysis above has been done with the assumption that we have a full duty cycle (Δ = 100%). For lower duty cycles, longer inactive periods will be inserted after a data subcycle. This will of course affect the sleep ratio in a positive way. The sleep ratio defined in (6.4) can than be written as:

$$\rho_i \;=\; \frac{T_{Dutycycle} - T_{on,i}}{T_{Dutycycle}} \;=\; \frac{T_{cycle} - (\Delta \cdot T_{on,i})}{T_{cycle}} \qquad (6.13)$$

Figure 6.8 shows the sleep ratio for different duty cycles. It can be seen that changing the duty cycle seriously affects the lifetime of the network. By setting a duty cycle of 50%, the sleep ratio is increased with 10%. When the duty cycle is 10%, the sleep ratio is higher than 99%. Of course, the value that can be used for the duty cycle depends on the amount of traffic in the WBAN. For low loads, a low duty cycle can be chosen, allowing a more energy efficient use of the network resources.

6.2.2 Throughput Efficiency

As in WASP, the throughput efficiency is defined as the percentage of useful traffic that can be sent over the network. It is the ratio of the number of slots during which data is sent to the sink and the length of a cycle. It thus expresses the amount of data traffic sent to the sink during 1 cycle. In this section, we determine the maximal throughput efficiency of CICADA, thus a duty cycle of 100% is assumed. The cycles follow directly one after the other. The throughput efficiency is then calculated as follows:

$$TPE = \frac{\sum_{i \in Ch_S} \alpha_i}{T_{cycle}} \tag{6.14}$$

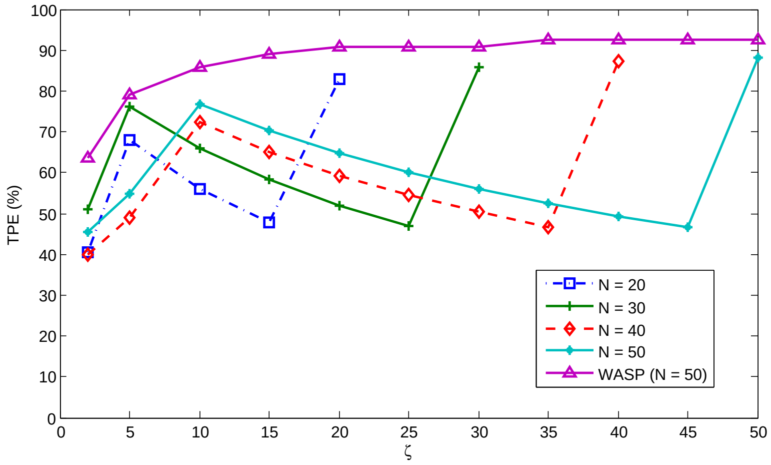
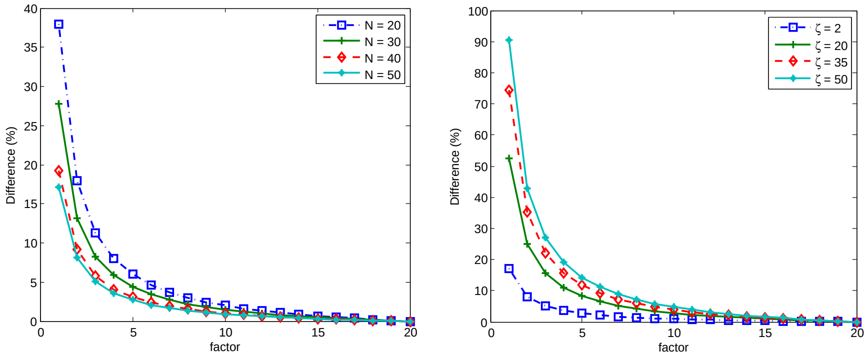


Figure 6.9: Throughput in CICADA for varying ζ and varying network size N. The slots in the control cycle are a factor 5 smaller (ϵ = 5).

The throughput efficiency (TPE) is shown in Figure 6.9 for a ζ-balanced tree and under the assumption that all nodes send one packet each cycle. As a consequence, the numerator simplifies to $N - 1$. It can be seen that the TPE is smaller than in WASP. This is due to the fact that all data is sent from the originating node to the sink in one cycle leading to longer cycles. The throughput efficiency also drops for some values of ζ. This can be explained as follows: if we look at the graph for $N = 50$, than the number of levels L is 3 for ζ varying from 5 to 45. When ζ increases, the denominator (which is the same as the denominator of (6.12)) increases as the control cycle length and the waiting period becomes larger. So the TPE will drop. For $\zeta = 50$, the number of levels is reduced to 2 and the cycle length will drop. It can be concluded that the number of levels should be kept small. The variation with respect to the size of the control cycle is limited and is omitted.



(a) Varying number of nodes. $\zeta = 5$

(b) Varying ζ. $N = 50$

Figure 6.10: Differences in throughput for varying ratio of the slot length in the control subcycle and data subcycle (factor = ϵ) compared with a ratio of 20. Full duty cycle is assumed.

Figure 6.10 shows the influence of the size of the control slot length. The difference in terms of percentage between the varying ϵ and $\epsilon = 20$ are plotted. For example, in Figure 6.10(a) for $N = 50$ and $\epsilon = 2$, a difference of 80% is found. This means that the throughput rises 80% when factor 20 is used instead of factor 2. In general, when ϵ is lower than five, a high decrease of the throughput is found. When the factor is ten or higher, almost no difference is seen. As a consequence, it is proposed to use an ϵ of at least ten in order to have a higher throughput.

The preceding formula gives the percentage of time in which data is received by the sink. It, however, does not give the efficiency of the whole system. Indeed, due to the tree structure of CICADA, simultaneous transmission in separate branches of the tree is possible. Consequently the throughput efficiency will be a lot higher. This is not reflected in (6.14) as it only considers the data received by the sink in one cycle. The total throughput efficiency is calculated as follows:

$$TPE_{total} = \frac{\sum_{i \in V_S} \alpha_i}{T_{cycle}} \tag{6.15}$$

The numerator gives all the slots in which data is sent, and strongly depends on the shape of the tree. For a ζ-balanced tree, the numerator can be written as

$$\sum_{y=1}^{L-1} \left(\zeta^y \cdot \sum_{i=0}^{L-y-1} \zeta^i \right) \tag{6.16}$$

Each level y has ζ^y nodes and each node on level y sends $\sum_{i=0}^{L-y-1} \zeta^i$ datapackets. However, this formula is only valid for a fully ζ-balanced tree. If the tree is not fully balanced, some branches of the lowest level will contain less than ζ siblings. We thus need to subtract these nodes $L - 1$ times as the data from these nodes is forwarded by each level. In the lowest level, there are $N - \sum_{i=0}^{L-2} \zeta^i$ nodes, whereas (6.16) supposes ζ^{L-1} nodes. As as consequence, the following needs to be subtracted from (6.16)

$$(L-1) \left(\zeta^{L-1} - N + \sum_{i=0}^{L-2} \zeta^i \right). \tag{6.17}$$

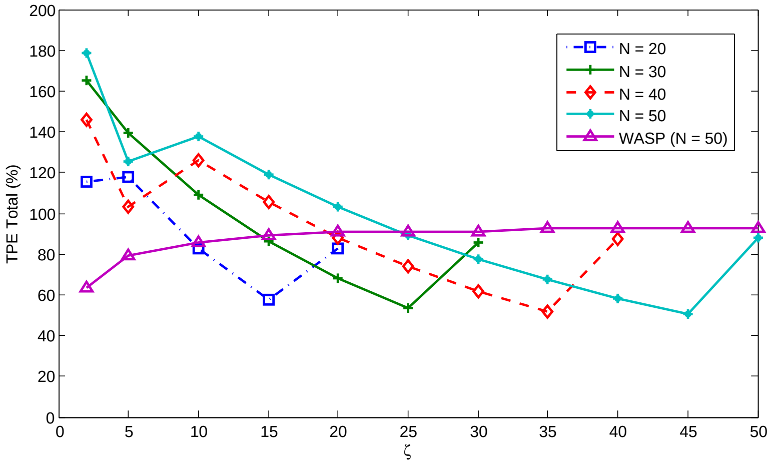


Figure 6.11: Total throughput efficiency in CICADA for varying ζ and varying network size N. The slots in the control cycle are a factor 5 smaller (ϵ = 5).

Figure 6.11 shows the total throughput efficiency. It can be seen that this efficiency is highest when ζ is small. This is to be expected as there will be more levels and it is more possible to send simultaneously. For example, when ζ equals 2, the tree breaks down in two branches that can sent simultaneously. Only when data is sent from the first level to the sink simultaneous transmission is not possible. Doing so, for large networks a TPE_{total} of almost 180% is achieved. When the number of levels is constant (e.g. for ζ = 10 until 35 for N =40), TPE^{total} drops as more nodes are on the same level end sending simultaneously is less possible.

For varying duty cycles, TPE and TPE_{total} scale with a factor Δ. Thus, for a duty cycle of 10$\%$, the throughput will be 10 times lower. Duty cycles can thus only be used in networks with small loads.

6.2.3 Mobility Support and Delay

Mobility support is a necessity for multi-hop Body Area Networks. A first cause is that the wireless transceivers have a short transmission range due to the low transmit power to save energy and due to the large attenuation. This results in very small scale topologies where even small movements can influence the link between two devices. Further, nodes attached to the limb will move relatively with respect to nodes attached on the rest of the body. Currently there is no mobility model available for WBANs. In our studies of the protocol performance we looked at simple, humanly feasible movements to study the impact.

CICADA supports mobility by the following mechanisms. (1) Nodes can join an existing network in 1 cycle: the moving node hears the new parent in its control subcycle and joins in its contention slot in the data subcycle. In the next cycle the node is assigned data slots and it can start sending. (2) A parent can monitor messages it expects to come from its children and vice-versa. Simulations show that marking a node as lost after 2 missed control packets suffices, as waiting for more cycles causes slow responses to a changed topology while tree stability is not improved much. (3) If a node knows it will move, it can send a DISCONNECT-message to its parent. These simple mechanisms make mobility support possible in CICADA: noticing parent loss and joining the new parent will take at most 3 cycles. Given the fact that cycles are short this results in good mobility support.

In TDMA protocols delay has an upper bound. As explained, CICADA uses a tree structure to send the data to the sink. Furthermore, the allocation of the different slots is done in such a way that in one cycle all data can be sent to the sink. CICADA allows a flexible increase or decrease of the number of slots assigned to a node after just one cycle, represented by δ_i in (6.1). This guarantees reasonably low delays even with variable bitrate traffic. A comprehensive study about the delay can be found in [5].

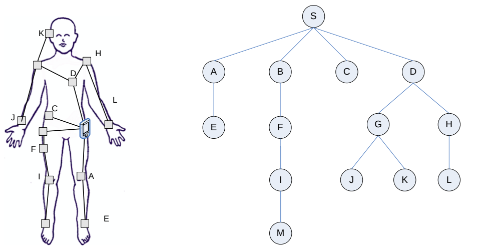
6.3 Implementation and Validation

6.3.1 Implementation

Like WASP, we have implemented the protocol in nsclick [6]. In order to take the specific properties of communication near the human body into account, we have used the advanced propagation model described in Section 2.3.1.3. Using this path loss on the human body, we have a more realistic view on the losses experienced near the human body and the corresponding influences on the network topology.

Further, we have used the energy model for the radio with the parameters of the Nordic transceiver, see Section 2.3.3.2. The bit rate is thus set to 1 Mbps.

6.3.2 Example Scenario



(a) Network on the body.

(b) Abstract view of the network.

Figure 6.12: Network topology used in the simulation

In order to allow a comparison between WASP and CICADA, we have taken the same example scenario. Figure 6.12(a) shows the example network where some sensors are placed on a human body. Figure 6.12(b) shows the generic tree view of this network. Thirteen nodes are each sending a CBR-stream to one sink with radios capable of transmitting up to 1Mbps. When configuring CICADA with control slots of 0.5 ms and data slots of 5 ms, the tree was set up after 192 ms.

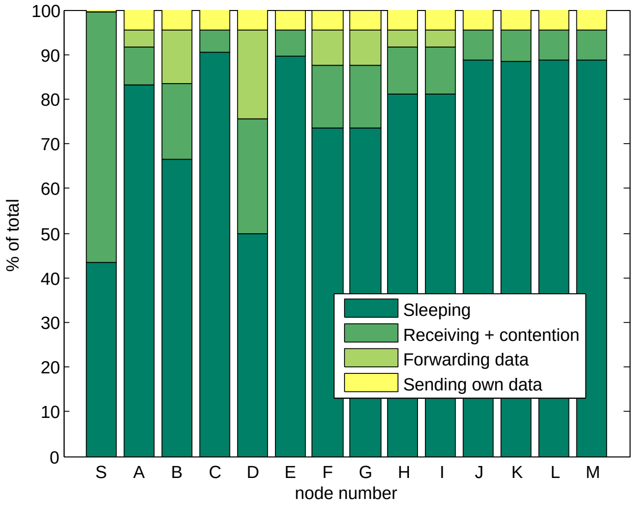


Figure 6.13: Time usage in the nodes. The sink has number 1, node A is number 2 and so on.

In the first evaluation, the duty cycle was set to 100% and all sensor data was sent to the sink. Each node sent a data packet of 500 bytes every 150 ms. When data was being generated at each node, the cycle length stabilized to 124.5 ms or 9 control slots and 24 data slots. For this example, the packet inter arrival time was set to be higher than the cycle length, ensuring that at most 1 packet was sent during each cycle. Doing so, no packet loss was obtained. As in WASP, nodes only have to operate their radio when they are sending or receiving data. The time usage of the nodes in CICADA is depicted in Figure 6.13. The sink can turn its radio off when it has sent its SCHEME-message in the control cycle (8 control slots) and in the waiting period in the data cycle (10 data slots). Thus, each cycle the radio can be turned off during 54 ms or 43.7% of the time. The same holds for the other nodes. The sleep ratios in this figure correspond to sleep ratios calculated with (6.4) and with the understanding that the radio is turned off when it has sent its control scheme. If we compare the energy usage of CICADA with WASP, Figure 5.9, it can be seen that for all nodes the sleep ratio is larger for CICADA. This is in line with the conclusions of Figure 6.7 where for low ζ the sleep ratio is higher for CICADA. In this scenario ζ corresponds to 4.

We have also calculated the overhead associated with CICADA, i.e. both the control packets as the headers of the data packets. The overhead for each node can be found in Table 6.5. Overall, the overhead is very limited, only a few percent. This was expected as the header length is limited (7 bytes for a data packet, see Section 6.1.11) and the control packets for sending the schemes are small. We also have to take in mind that in this example the size of the payload is 500 bytes. The overhead of the sink is 100% as the sink only sends control packets.

Node	Total bytes sent	Overhead (bytes)	Ratio
S	1674	1674	100.00%
A	47800	1800	3.85%
B	94640	2640	2.85%
C	24231	1231	5.19%
D	141499	3499	2.53%
E	24180	1180	4.99%
F	71157	2157	3.10%
G	71157	2157	4.62%
H	70741	1741	2.50%
I	47741	1741	3.73%
J	24180	1180	4.99%
K	24180	1180	4.99%
L	24180	1180	4.99%
M	24180	1180	4.99%

Table 6.5: Overhead induced by CICADA for the example network of Figure 6.12(b)

The example network was also used to evaluate the sleep ratio and packet loss for varying packet inter arrival time ($t_{arrival}$) and for different duty cycles. The packet inter arrival time is defined as the time between two packets received from the application layer. CICADA is also compared with S-MAC [7], a much used CSMA-style protocol for sensor networks. Nodes synchronize on time by building virtual clusters and employ a fixed duty cycle to reduce idle listening overhead. S-MAC includes carrier sense, collision avoidance (RTS/CTS signaling), and overhearing avoidance. We have used the implementation of S-MAC available in NS-2 with the default parameters (a duty cycle of 10%) and with fixed (optimal) routing. Doing so, the overhead and delay caused by the routing protocol is avoided. In Figure 6.14, it can be seen that CICADA has a packet loss of 0% as long as the packet inter arrival time is lower than the length of the duty cycle. The main reason for this problem is that the node is not aware of the application's traffic pattern (i.e. the sensor rate). The node can only reserve slots (by changing its α_i) for data already queued. For example, when the nodes start transmitting packets, two packets will be reserved in the first cycle. The cycle lenght will increase and in the next cycle three packets need to be reserved. This will again lenghten the duration of the cycle, and the following cycle four slots will be reserved. This mechanism creates an avalanche effect in the length of the duty cycle. Currently, CICADA does not take this problem into account. A possible solution is to implement a traffic predictor who can more precisely predict the packet inter arrival time. Compared with S-MAC, CICADA performs a lot better.

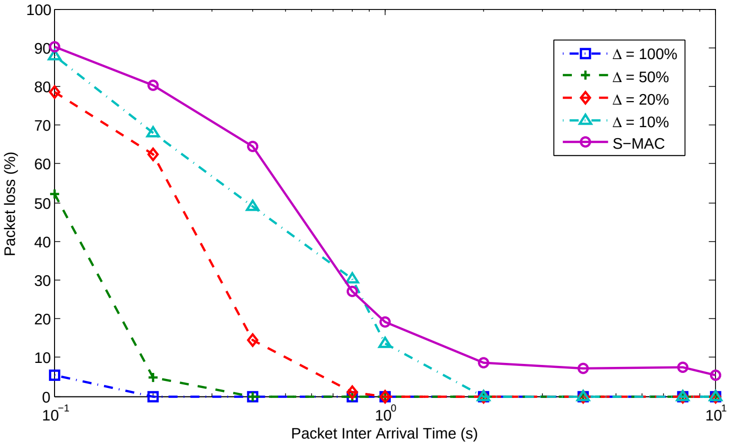


Figure 6.14: Packet loss for the network of Figure 6.12(b). The duty cycle Δ and the packet inter arrival time $t_{arrival}$ is varied.

The sleep ratio is shown in Figure 6.15. The sleep ratio drops for small packet inter arrival times, which is expected as more traffic is sent in the network. However, when the packet inter arrival time is too low, the sleep ratio rises again. More packets are lost due to buffer overflows, thus less packets are actually sent. This effect further scales with the packet loss of Figure 6.14. For lower duty cycles, the sleep ratio rises up to 95%. Compared with S-MAC (duty cycle of 10%) CICADA performs better when the duty cycle is set to 20% or lower.

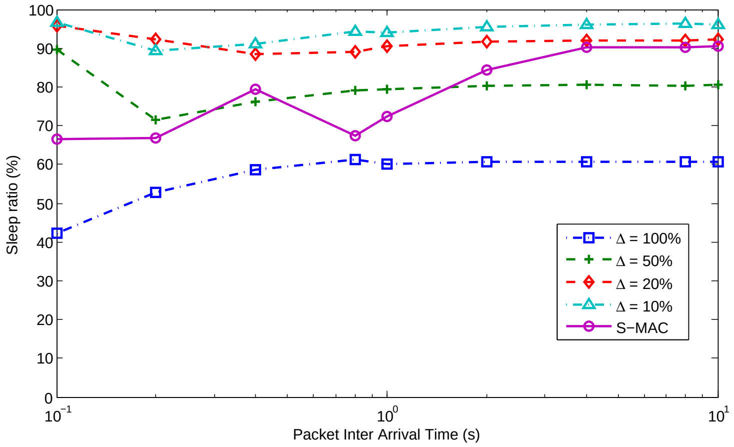


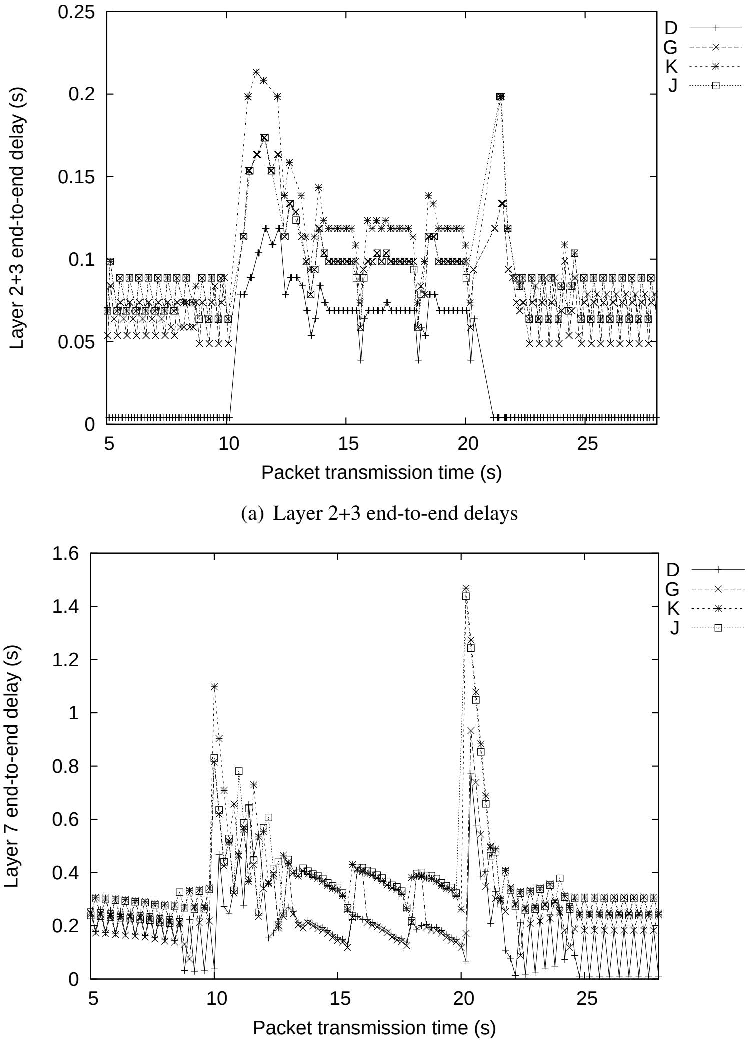
Figure 6.15: Sleep ratio for the network of Figure 6.12(b). The duty cycle Δ and the packet inter arrival time is varied.

The results above show that CICADA has a high sleep ratio, especially when the duty cycle is reduced to 50% or lower. Longer duty cycles however lead to packet loss when packet inter arrival times are too low. When deploying a WBAN using CICADA, one thus has to balance the desired throughput and the energy efficiency against each other.

6.3.3 Example Scenario: Mobility

Two scenarios where a subset of the nodes move are considered. In the first scenario, node D moves after 10 seconds from its original position toward node C, drastically changing the topology. Ten seconds later, node D is brought back to its original position. In the second scenario, node J moves closer to the sink. This corresponds with the movement of the right arm.

Figure 6.16(a) shows the layer 2+3 end-to-end delays of nodes D, G, K and J for the first scenario. Before the movement starts, the delay is between 0.05 and 0.1 ms, node D almost experiences no delay. These delays are very low due to the ability to deliver all data packets to the sink in one cycle. Between 10s and 20s, node D joins with node C. This influences all the nodes below node D as an extra hop is required to reach the sink. For example, the data from node G



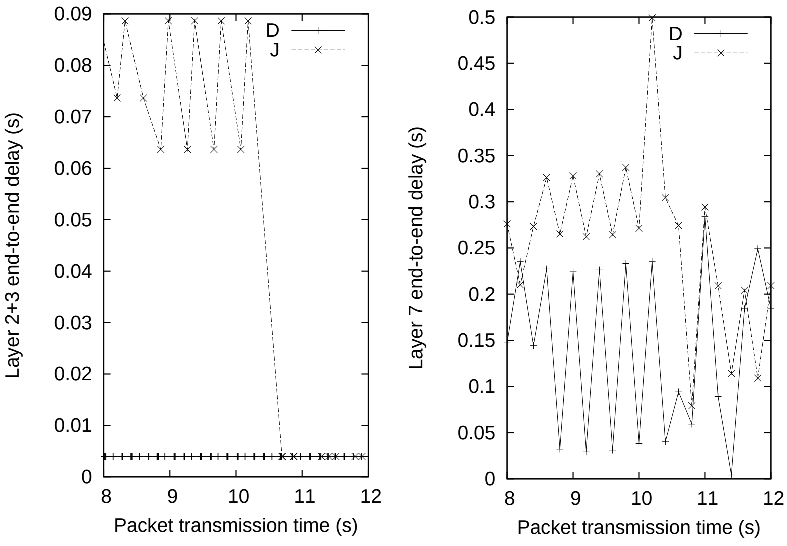
(a) Layer 2+3 end-to-end delays

(b) Layer 7 end-to-end delays

Figure 6.16: End-to-end delays on different layers in the example scenario of Figure 6.12(b). At 10s, node D is moved toward node C and brought back at t = 20s.

now follows the following path: G–D–C–S. Hence the delay increases for all the depicted nodes. When node D moves, a transitional phase is noticed with larger delays. These are caused by the fact that the node waits 2 cycles before searching a new parent. As a node releases all of its children when it has lost his parent, all the nodes below node D need to rejoin. This causes an additional delay. The delay is the highest for node K. While the nodes are disconnected, data is buffered in the nodes. When the nodes rejoin the network, they empty their data buffers. This lengthens the cycles, increasing the delay. The combination of joining and buffering effects explains the large transitional delay. This phase lasts about two seconds. Notice that when the network connection is restored, delays are low again.

The delay at the application layer, i.e. layer 7 end-to-end delay, experiences the same evolution as the delay at layer 2+3, see Figure 6.16(b). The delay is higher as there is a mismatch between the moment the packet is delivered to the network layer and the moment the packet can be sent in the following cycle. As the cycle length is low, i.e. 110 ms in this scenario, the delay remains below 0.35 seconds in steady state, which is very low. Again, a transitional phase happens.



(a) Layer 2+3 end-to-end delays

(b) Layer 7 end-to-end delays

Figure 6.17: End-to-end delays on different layers in a mobile network. Node J is moved closer to the sink.

The results of the second scenario are shown in figures 6.17(a) and 6.17(b). At 10s, node J chooses the sink as parent. Consequently, the layer 2+3 delay for node J decreases. Looking at the application delay, we once again see a small transitional phase with higher delay.

These small scenarios show that CICADA is capable of quickly coping with a changing topology.

6.3.4 Real World Implementation

As part of the IBBT-IM3 project, CICADA is currently being implemented[2] on an existing transceiver, the NxH1001 CoolFlux transceiver. This receiver uses magnetic induction for the communication, offers 300 kbps raw bandwidth and support for TDMA channel access. The transceiver is based on the CoolFlux DSP, an ultra low power C-programmable DSP core [8].

Preliminary results with three radios show that the protocol works properly. The following setting was used to evaluate the protocol. One of the radios is used as personal device, the two other as sensor devices (one measuring the temperature and the other an ECG). In a first step, both radios are directly connected to the sink. All data is correctly received by the sink. Then, one of the radios is moved outside the range of the sink, but within the range of the other sensor. Within a second, the data of the device is relayed by the other sensor towards the sink.

6.4 Increasing the Reliability

From the previous examples, it was clear that CICADA experiences a low packet loss. Although, in a more dynamic environment, a higher packet loss can occur due to interference or the way the tree is set up. Therefore, we will add more reliability mechanisms to the protocol. Three adaptations are envisioned: the introduction of an acknowledgment map, randomization of the schemes and repeating the SCHEME-messages of the parent (aka overhearing).

6.4.1 Acknowledgments

The way acknowledgments are used is similar to that of WASP (see Section 5.3.1.6). At the end of each SCHEME-message, an ACK-sequence for the previous cycle is sent: it contains a bit for each slot in which data was sent. A 0 denotes that the packet was not received correctly, a 1 denotes success. The node the SCHEME-message is destined for, say node i, will check the ACK-sequence. If the position of a 0 corresponds to one of the slots where i sent data, the node will resend that data in this cycle. Its parent will already include an extra slot for node i by adjusting it's α_i and β_i. For example, in the network of Figure 6.1 the sink S adds the following ACK-sequence to it's scheme: 11101. Nodes A and B receive these messages and node B knows that it has sent in the fourth data slot. Node B will retransmit that packet and delete the other packets from its buffer.

[2] The implementation is mainly done by Wim Torfs of the University of Antwerp

6.4.2 Randomization of the Schemes

By randomization of the schemes, we try to restrict the influence of interference. When the tree is set up, it might happen that two nodes i and j can hear each other, but have different parents, k and l respectively, e.g. because the link between k and i is more reliable than the link between l and i. When the schemes are fixed, i.e. the first node to join always sends first and so on, it might well happen that nodes i and j will interfere while sending their data to their respective parents. By randomizing the schemes, i.e. by changing the sequence in which the children are allowed to send, the overall interference will be decreased.

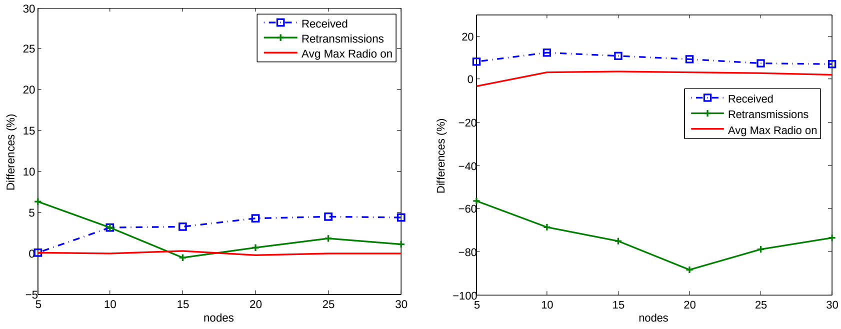
6.4.3 Overhearing

During our simulations we noticed that, from time to time, nodes miss a SCHEME-message from their parent, because of a link that's not very stable. The result is that this node and all nodes below it cannot do anything and must have their radio on until the next cycle. In order to tackle this problem, a child node repeats the scheme of its parent when it sends its own scheme. Doing so, siblings can exploit this information if they missed it.

6.4.4 Analysis

The path loss model (2.4), the link probability (2.12) and the improvements are implemented. The simulator was used to analyze the changes to the protocol. The size of the network is varied from 5 to 30 nodes and the nodes were randomly placed in a 2 by 2 meter area with the sink positioned in the center. The distances between the nodes is at most 40 cm in a connected topology, i.e. every node is within transmission range of at least one other node so there is always a path to the sink. Nodes start randomly, they do not join the network all at once. All simulations were run during 10.000 slots for 1000 randomly generated topologies, while making sure the same topologies are used in comparisons. Each node generates one packet after a fixed period. This period is defined in slots and equals three times the number of nodes in the network. The number of packets received by the sink and the number of retransmissions are considered. We also look at the number of slots the radio was on, averaged out over the number of runs, to study the impact on network lifetime. The randomization and overhearing mechanism are evaluated separately and combined. The ACK-mechanism is always used.

Results of the simulations are shown in Figure 6.18(a). The values represent the improvement in percentage between the results without and with randomization. It can be seen that scheme randomization has a positive impact. The number of packets that can be received by the sink increases by more than 4% for larger networks. This is caused by the lower number of collisions. Yet, it can be seen that



(a) Randomization

(b) Overhearing

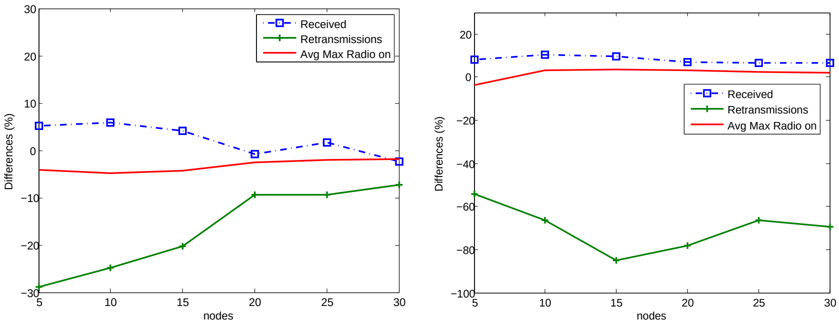
Figure 6.18: Evaluation of the proposed reliability mechanisms. The figure plots the difference between regular CICADA and CICADA with the proposed reliability mechanism. For example, in (a) the reliability mechanism has 5% more received packets than regular CICADA for a network of 20 nodes.

the number of retransmissions is larger. The rise of the number of retransmissions, however, is lower than the rise of the number of received packets. Thus, relatively spoken, the number of retransmissions has not increased. The absolute increase can be explained by the higher number of transmissions in the network. For small networks, little effect was found as the parent nodes have few child nodes to randomize. It is important to notice that the figure also shows that the average time a radio is on is almost similar with or without randomization. This is expected as the number of slots a node sleeps does not change. Hence, the scheme randomization leads to an higher throughput in the system while having almost no impact on the network lifetime.

The results for the overhearing mechanism are shown in Figure 6.18(b). We see that the sink receives about 10% more data. The chance of missing the parent's scheme is lower as the scheme can be recovered by listening to the sibling's control packets. The node now knows when to send its data, which will increase the throughput. The number of retransmissions drops dramatically with roughly 70-80%. This shows that overhearing has a positive effect as less control messages are missed and the nodes thus know when to send their data, leading to fewer retransmissions. The impact on the energy consumption is very low.

The overhearing solution also increases the overhead. When the scheme of the parent is included to the SCHEME-message, an additional of $4 + 2 \cdot x_p$ bytes are added, where x_p indicates the number of children of the parent. If we assume that in a network each node has a maximum of 10 children, the length of the control packet will change from 25 bytes to 49 bytes. This means that the length of the slot size in the control cycle needs to be increased, which will have an impact on the energy efficiency. However, these influences are minor as long as the slot ratio

ϵ is larger than ten, as was shown in Section 6.2.2. Hence, if a data slot can hold a message of 500 bytes, the influence of adding your parent's scheme is minor.



(a) Overhearing with low packet interarrival time

(b) Combined solutions

Figure 6.19: Evaluation of the proposed reliability mechanisms (2).

We have also evaluated the influence of a low packet inter arrival time. Each node now sends data after a period of 1.5 times the number of nodes in the network. For smaller networks, this is higher than the duration of the cycle, for larger networks not. Results can be seen in Figure 6.19(a). The number of retransmissions caused by the overhearing mechanism drops for larger networks. The network is already congested and the overhearing mechanism only gives a limited solution for a congested network.

The combined effect of these two mechanisms are shown in Figure 6.19(b) for high packet inter arrival time. It can be seen that the effects of the overhearing mechanism dominates. The combined solution further has little influence on the energy performance.

6.5 Improved Parent Selection

In the current implementation of CICADA, a node joining the network listens for the SCHEME-messages and sends a JOIN-REQUEST message to the node which the first received SCHEME-message was from. This might not always be the best parent, as the network topology can change over time due to nodes that disappear or are added to the network. Further, by choosing another parent, a more optimal network topology can be obtained, as discussed in Chapter 3.

The parent selection can be improved by letting the node wait until packets from several nodes have been received. Using this information, the most optimal or most reliable node can be selected as a parent. For example, a node that is attached to a constantly moving arm, can choose the node with the most reliable connection as a parent, e.g. a node attached to the shoulder.

In the control subcycle, the node always hears schemes from nodes at a higher level. Now, in addition, it will also listen to the scheme of siblings, for a predetermined number of cycles. When a broadcast is received from node j, the node is added to a parentlist. The node only considers the broadcast of nodes at the same level or a higher level. Doing so, the node cannot join any of its children and loops are avoided. Further, the nodes can stop listening sooner, which will be beneficial for the energy consumption. After a few cycles, the node can determine its most optimal parent, based on some predefined metric such as load of the parent, link quality, available energy at the parent and so on. When the node sends a JOIN-REQUEST-message to his new parent, it also sends its current α and β. Doing so, the node keeps its children. This will lower the overall latency as the children do not need to reconnect.

As an extra improvement, a node can change its transmission power or use a more robust coding scheme at the PHY-level if it does not find a parent that meets the desired criteria. Doing so, the node will have more parents to choose from.

Finding optimal metrics for choosing a parent is an important next step in the development of CICADA.

6.6 Adding Security

The CICADA protocol, as it is described in the previous sections, does not guarantee any form of security. Unauthorized nodes can easily join the WBAN, and all communication in the network is sent in plain text and is not integrity protected. To counteract these problems, appropriate security mechanisms have to be added to the CICADA protocol. The result is the CICADA-S protocol, the secure version of the CICADA protocol. This protocol is described in Appendix C [3]. In this section, we will briefly discuss the security mechanisms. The security is handled by exchanging and updating keys.

From security point of view, there are four main states which take place during the lifetime of a sensor: the secure initialization phase, the sensor (re)joining the WBAN, a key update procedure in the WBAN and the sensor leaving the WBAN. It is assumed that the sink can set up a secure connection with a back-end security server, placed at a hospital or data center. In the initialization phase, each device gets a random generated secret key (k_A) from the security server. This is done by authorized personnel in order to guarantee the integrity of the key. In the (re)joining phase, the sensor device sends a JOIN-REQUEST message in the contention slot, secured with the secret key k_A. A counter CTR_A is stored in the memory. It is used to assure the freshness of the message and to avoid replay attack. Only messages with the correct counter value are considered. The

[3] This work was done in collaboration with Dave Singelée of the K.U. Leuven who specializes in the security aspects of wireless sensor networks.

JOIN-REQUEST message is forwarded to the sink. When the gateway receives the secure JOIN-REQUEST message of sensor *A*, it forwards this request to the back-end server via the secure end-to-end channel. This triggers a protocol in which the key k_A is securely transported from the back-end server to the gateway. From the moment the gateway has access to the key, it can check the validity of the JOIN-REQUEST by verifying the message authentication code, and in case of a rejoin, also the value of the counter CTR_A. If this is successful, the sensor is allowed to join the WBAN and is assigned a temporary identity $localID_A$. This temporary identity, which is chosen by the medical hub, is established in order to preserve the privacy. The node is also informed of the group key s which is used for all the normal communication and is shared by all the nodes of the WBAN. This group key is updated regularly as follows: first, the gateway randomly generates a new group key s. Next, it performs a secure key transport procedure with all the nodes in the WBAN. The gateway constructs a message, unique for every sensor, which contains the encrypted value of the updated group key s. These messages are broadcasted to the sensors in the control cycle. When a node leaves the network, a new group key will be generated. A discussion on the performance overhead and security claims can be found in Appendix C.

By adding these security mechanisms, CICADA becomes the first integrated solution that copes with threats that occur in medical monitoring scenarios. It is shown that the integration of key management and secure, privacy preserving communication techniques within the CICADA-S protocol has low impact on the power consumption and throughput.

6.7 Mapping of CICADA on MOFBAN

In this section, we will describe how the CICADA protocol can be used in the MOFBAN framework from Chapter 4. The functionality of CICADA can be split up in different modules. We will briefly describe which modules can be used and how the functionality is implemented.

- The transmission module sends the packets and is responsible for the time synchronization. It is invoked by the controller at the beginning of the slot where the node has to send.

- The routing module handles the control packets and is responsible for building the control and data schemes. When a control packet is received, the routing module extracts the control and data scheme and determines the slots in which the radio has to send data. This is sent to the controller who sets a timer. When the timer expires, the transmission module is invoked. Based on the ChildTable, the routing module calculates the data scheme for its children. A control packet is made and the timer is set to invoke the transmission module. It is also responsible for handling incoming data packets. It extracts the α and β-values and updates the ChildTable.

- The self organization module is used for building the tree. When a node joins the network, the controller automatically starts the self organization module. When a packet is received, it is forwarded to the self organization module. The module selects the parent, determines the time slot of the selected parent and creates a JOIN-REQUEST message. This message is stored by the controller in the information module. At the correct time, the controller invokes the transmission module which sends the JOIN-REQUEST message. There exist several types of this module. If a higher reliability is desired, another mechanism for parent selection can be used, as described in Section 6.4. This can be altered by simply plugging in a new module.

- The ACK-module implements the ACK-mechanism described in Section 6.4.1. When a message arrives at the node, it is first sent to this module and afterwards to the routing module. When no packet is received when one is expected, the module updates the ChildTable and updates the ACK-sequence. This module is optional.

- The security module is optional and implements the security mechanisms described in Section 6.6. When the routing module has made a data packet, it is first sent to the security module to encrypt the message. When a packet is received, it is first decrypted by the security module and then sent to the ACK-module or the routing module.

6.8 Conclusion

In this chapter we have presented CICADA, a cross-layer protocol for WBANs. Based on a data gathering tree structure, CICADA controls the communication in a WBAN using distributed slot assignment. By dividing the time into slots and by reserving the slots, collisions and overhearing is avoided, leading to low energy consumption. All the packets can reach the sink during only one cycle, severely reducing the delay. The use of a control subcycle as well as a data subcycle results in low packet loss and high sleep ratios while the network flexibility is preserved. In the control subcycle, each node determines the slots for the data subcycle in which its children can send, based on the assignment received from its parent. Further, data-aggregation and a duty cycle Δ is used to even further improve the lifetime of the network. By lowering the duty cycle, a higher sleep ratio is achieved, but the available throughput lowers inversely proportional.

We have evaluated the performance of CICADA in terms of energy efficiency, throughput, mobility support and delay. It was shown that for a high sleep ratio the tree should not have too many levels and that the energy efficiency of CICADA is better than WASP's, especially when the ratio of the lengths of the control subcycle slots and data subcycle slots is high enough. The throughput, however, is about 20% lower than CICADA's.

The protocol was implemented and validated in NS-2. It was shown that the packet loss is 0% when the packet inter arrival rate is higher than the length of a full cycle. The protocol performs considerably better than S-MAC, both in terms of packet loss and sleep ratio. Mobility is handled quite well.

The reliability of CICADA was evaluated and additional mechanisms were proposed in order to improve the reliability even further, such as the randomization of schemes and overhearing the control messages sent by the siblings. We have added security to CICADA, making it the first integrated solution that copes with threats and privacy issues. As a last part, it is described how CICADA can be implemented in the MOFBAN framework.

If desired, CICADA is compatible with legacy systems such as IEEE 802.15.4. A device can use the IEEE 802.15.4 PHY-layer and the unbeaconed MAC-layer without the use of acknowledgments. A small adaptation is needed for dividing the time axis into slots. CICADA can work on top of this structure.

CICADA can still be improved, such as the inclusion of a more improved slot synchronization, similar to the one used in [9] that provides synchronization in a tree network. Next to that, the packet loss can be further reduced when cycle length is higher than the packet inter arrival rate by adding a traffic predictor.

References

[1] S. Lindsey, C. Raghavendra, and K. M. Sivalingam. *Data gathering algorithms in sensor networks using energy metrics*. IEEE Transactions on Parallel and Distributed Systems, 13(9):924–935, September 2002.

[2] G. Lu, B. Krishnamachari, and C. S. Raghavendra. *An adaptive energy-efficient and low-latency MAC for data gathering in wireless sensor networks*. In Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International, April 2004.

[3] S. Gandham, Ying Zhang, and Qingfeng Huang. *Distributed Minimal Time Convergecast Scheduling in Wireless Sensor Networks*. 2006. ICDCS 2006. 26th IEEE International Conference on Distributed Computing Systems, pages 50–50, 2006.

[4] S. Upadhyayula and S. K. S. Gupta. *Spanning tree based algorithms for low latency and energy efficient data aggregation enhanced convergecast (DAC) in wireless sensor networks*. Ad Hoc Netw., 5(5):626–648, 2007.

[5] B. Latré, B.Braem, I.Moerman, C. Blondia, E. Reusens, W. Joseph, and P. Demeester. *A Low-Delay Protocol for Multihop Wireless Body Area Networks*. In Mobile and Ubiquitous Systems: Networking & Services, 2007 4th Annual International Conference on, Philadelphia, PA, USA, August 2007.

[6] M. Neufeld, A. Jain, and D. Grunwald. *Nsclick:: bridging network simulation and deployment*. In MSWiM '02: Proceedings of the 5th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems, pages 74–81, New York, NY, USA, 2002. ACM.

[7] W. Ye, J. Heidemann, and D. Estrin. *Medium Access Control with Coordinated, Adaptive Sleeping for Wireless Sensor Networks*. IEEE/ACM Trans. on Networking, 12(3):493–506, June 2004.

[8] Coolflux DSP [online] http://www.coolfluxdsp.com/.

[9] L. Dai, P. Basu, and J. Redi. *An Energy Efficient and Accurate Slot Synchronization Scheme for Wireless Sensor Networks*. In Broadband Communications, Networks and Systems, 2006. BROADNETS 2006. 3rd International Conference on, pages 1–8, San Jose, CA, USA,, October 2006.

7
Overall Conclusions

During this research work, we focused on the development of energy efficient and reliable network protocols for wireless Body Area Networks. As the research in this domain is fairly new, we have also considered different aspects of networking in a WBAN. This dissertation therefore contributed to the following areas of research:

- Based on existing studies of communication around the human body, a probabilistic connectivity model and an energy model of the radio is proposed;
- A thorough study of possible network topologies in terms of energy efficiency and reliability in a WBAN was performed by means of a line and tree topology and an ILP-formulation for more general networks;
- A modular framework for WBANs was proposed: MOFBAN;
- The IEEE 802.15.4 protocol was analyzed and its suitability for WBANs was verified;
- Two cross layer protocols for WBANs where developed and analyzed: WASP and CICADA.

The taxonomy and the special requirements of a WBAN were briefly discussed in Chapter 2. It was shown that a WBAN imposes specific challenges on the network in terms of energy efficiency, reliability and heterogeneity. The dissimilarities between a WBAN and sensor networks were explained. Differences include the scale of the network, even more rigorous restrictions on the energy consumption and required reliability, mobility support and heterogeneity. The taxonomy indicated that the electromagnetic waves along the body experience high path losses. Especially when the devices are not in line-of-sight, a path loss exponent of almost 6 is found. When it turned out that no connectivity model existed that takes this into account, a new model was formulated. This model determines the probability of a connection based on the distance between the sender and the receiver. The model uses a lognormal distribution for determining the range of the node instead of assuming a circular coverage. Based on an energy model for the radios in WSNs, a new energy model was presented and the parameters for three frequently used radios were determined.

Once the physical layer of the WBAN was investigated, the connectivity model and the energy model were used for determining possible network topologies for a WBAN. In particular, we focused on two major characteristics: the energy efficiency and the reliability. For the former, we have defined the lifetime of the network as the time for the first node to die. A basic comparison in a line and tree topology showed that multi-hop communication performs significantly better for nodes far away from the personal device, but not for nearer nodes. This has led to the introduction of cooperation in the network where the sensor nodes cooperate in delivering the data. The farthest nodes send their data to nearer nodes that forward the data directly to the sink, together with the self-generated data. An ILP-formulation was used to evaluate the effect of cooperation and indicated, on average, a lifetime increase of more than 20%. As a next step, relay devices were added to the network, further improving the lifetime of the network. By combining cooperation and the use of relay devices, an even more energy efficient network was obtained. An additional lifetime increase up to 30% was found. When data-aggregation is used, an even higher gain can be achieved. The reliability of the network was subsequently added to the ILP-formulation by using the previously defined connection probability. By choosing more reliable links while giving consideration to the energy consumption, the reliability can be increased with only a very limited impact on the lifetime of the network. Overall, it was concluded that a multi-hop architecture is the best choice for a WBAN and that one has to deal with a trade-off between energy efficiency and reliability. Adding relay devices is helpful for both the energy efficiency and reliability.

Based on the taxonomy, we realized that, in order to support the complex heterogeneity in a WBAN, the current solutions for network architectures are not effective. Instead of choosing the most suitable protocol for each network layer and selecting the correct settings for each different application, new and more efficient approaches were needed. Hence we proposed MOFBAN, a lightweight modular framework for WBANs. Due to its modular structure and its API, it is more transparent for the application designer and easily adaptable and expandable. Duplication of functionality is avoided and heterogeneity is well supported by using different implementations of the modules. Depending on the capabilities of the node and/or the application, more modules (and thus network functionality) can be added. The interaction between the modules is handled by the controller module and information is stored in a common database. It further is responsible for activating the appropriate module at the correct time by using the scheduler, integrated in the controller module. The controller module can be considered as the heart of MOFBAN. To enable communication between the nodes and the modules, a modular header structure was presented.

Once the network architecture was defined, the existing MAC-strategies for WSNs and WBANs were investigated. More specifically, the maximum throughput and minimum delay of IEEE 802.15.4 were determined. We have expressed the exact formulae for calculating the maximum throughput of the unbeaconed version of IEEE 802.15.4. It was concluded that the throughput scales with the packet size and that in the 2.4 GHz band a maximum throughput of 163 kbps or an efficiency of 64.9% can be achieved. The bandwidth efficiency is rather low due to the small packet size, imposed in the standard. Further, it was argumented that IEEE 802.15.4 is not the best solution for supporting communication in WBANs because of the high energy consumption and overhead. A short analysis of existing MAC and routing protocols showed the lack of an integrated and energy efficient solution. As a consequence, we have presented two new network protocols that handle both the channel medium access and the routing aspects.

The first network protocol is called WASP. It sets up a spanning tree in a distributed manner and uses timeslots. Every node sends out a proprietary WASP-scheme to inform the nodes of the following level when they are allowed to send. These WASP-schemes are generated locally in each node. It is shown that the throughput efficiency can reach up to 94%, depending on the number of levels used. The end-to-end delay is shown to be low, fixed and related to the number of levels in the tree. WASP does not support bidirectional traffic. WASP was implemented and validated in NS-2. WASP offers a first answer to the challenges encountered in a WBAN. Still, several improvements can be made in terms of energy efficiency and reliability.

As a last part, CICADA has been presented. It uses a data gathering tree and controls the communication using distributed slot assignment. The use of a control subcycle as well as a data subcycle results in low packet loss and high sleep ratios while the network flexibility is preserved. It also enables two-way communication. Data-aggregation and the use of a duty cycle Δ even further improved the lifetime of the network. By lowering the duty cycle, a higher sleep ratio is achieved, but the available throughput lowers inversely proportional. Evaluation of CICADA showed that for a high energy efficiency the tree should not have too many levels. The packet loss of CICADA is 0% when the packet inter arrival time is higher than the length of a full cycle. The reliability was evaluated and additional mechanisms were proposed in order to improve the reliability even further, such as the randomization of schemes and overhearing the control messages sent by the siblings. A security mechanism was added to CICADA. Doing so, CICADA became the first integrated solution that copes with threats that occur in medical monitoring scenarios. It was shown that the integration of key management and secure, privacy preserving communication techniques within the CICADA-S protocol has low impact on the power consumption and throughput.

In the future, we expect work to continue on extensions of some of the key aspects discussed in this dissertation. The energy model presented in this thesis can be improved by adding the cost of wake-up and the turnaround time between sending and receiving. Further, the development of new ultra low power radios will lower the energy consumption even further and can have an influence on the network topology. For example, a lot of effort is put in the development of radios using UWB. The use of these radios will significantly alter the communication along the human body, leading to a new energy model for the radios.

A second notable extension can be found in the development of the modular framework. The proposed framework needs to be properly validated and the impact of the framework on the overall performance of the network should be thoroughly investigated. The proposed concepts, such as the modular header structure, needs to be investigated further. The ideas are not only useful for a WBAN, but can also be used as a starting point for a more complex framework that can be used in wireless sensor networks. Further, more QoS-mechanisms can be added to the framework. This is already subject of ongoing research in the IBCN research group.

Finally, CICADA can still be improved. For example, more effort should be put in maintaining the slot synchronization. This can be done using the tree structure of the network or via synchronization with the heartbeat, as was done in H-MAC. Further, the problem of the experienced packet loss when the cycle length is higher than the packet inter arrival time should be tackled by for example adding a traffic predictor. CICADA can further be adapted to take into account the heating of devices. When the temperature of a device becomes too high, the device is no longer used as relay station, i.e. it no longer uses a contention slot. The packets are thus withdrawn from the heated zones and rerouted through alternate paths. If needed, extra relay devices can be added to the network. More focus can be given to the interaction with the physical layer, e.g. by changing the transmission power. Further, the improvement of QoS to manage and reserve communication resources, such as bandwidth specification and reservation. The aggregated throughput can be improved by letting the nodes at the lowest level start sending while the nodes on the highest level are still transmitting. Doing so, links that have no probability of interference can speak at the same time. The impact on the delay needs to be investigated further. CICADA further only offers a limited answer to moving nodes, i.e. nodes attached to a moving limb which causes frequent topology changes. Improved mobility handling should be added to CICADA. Bart Braem of the PATS research group at the University of Antwerp is already working on this. In this respect, it is also interesting to have a good mobility model for nodes in a WBAN. For making such a mobility model, it is necessary to investigate further the propagation round the human body, in rest and in motion.

A WBAN is expected to be a very useful technology with potential to offer a wide range of benefits to patients, medical personnel and society through continuous monitoring and early detection of possible problems. With the current technological evolution, sensors and radios will soon be applied as skin patches. Doing so, the sensors will seamlessly be integrated in a WBAN. Step by step, these evolutions will bring us closer to a fully operational WBAN that acts as an enabler for improving the Quality of Life. We feel that this dissertation has contributed in this evolution and should certainly not be considered as an endpoint, but as a source of inspiration for future research directions.

A

Throughput and Delay Analysis of Unslotted IEEE 802.15.4

B. Latré[1] P. De Mil[1], I. Moerman[1], B. Dhoedt[1], N. Van Dierdonck[2], P. Demeester[1]

published in The Journal of Networks, 2006, Vol. 1, No. 1, pp. 20–28

[1] Ghent University – IBBT, Dept. of Information Technology, IBCN research group, Gent, Belgium

[2] Ubiwave N.V., Zele, Belgium

Abstract

The IEEE 802.15.4 standard is designed as a low power and low data rate protocol offering high reliability. It defines a beaconed and unbeaconed version. In this work, we analyze the maximum throughput and minimum delay of the unbeaconed or unslotted version of the protocol. First, the most important features are described. Then the exact formula for the throughput and delay of a direct transmission between one sender and one receiver is given. This is done for the different frequency ranges and address structures used in IEEE 802.15.4. The analysis is limited to the unslotted version as this one experiences the lowest overhead. It is shown that the maximum throughput depends on the packet size. In the 2.4 GHz band, a bandwidth efficiency of 64.9% is reached when the maximum packet size is used. Further we describe the influence of the back off interval. A significant *gain is found when the backs off parameters are altered. We have measured the throughput experimentally in order to compare the theoretical analysis with real-life examples.*

A.1 Introduction

In the last few years, new applications using different kinds of wireless technology are rapidly emerging. These applications all have their proprietary requirements with regard to the required data rate, power consumption, reliability and much more. Hence, several new protocols have been proposed such as IEEE 802.11g, IEEE 802.16 and IEEE 802.15.4. Whereas the first two focus on achieving higher data rates in order to support high bit rate applications, the latter is designed for low data rate and provides high reliability for activities such as controlling and monitoring. These applications generally use simple devices, such as sensors, which are not capable of handling complex protocols as they have limited processing capacities and limited power available. An interesting application is building automation. The building is equipped with wireless devices such as temperature sensors and light switches. These devices are battery operated and consequently require a low power protocol. In addition, the data that needs to be sent over the network is limited to a few kilo bits per second. More application scenarios are defined by the ZigBee Alliance [1] that defines the routing and application layers above the IEEE 802.15.4 standard. The goal of the IEEE 802.15.4 standard is to provide a low-power, low-cost and highly reliable protocol for wireless connectivity among inexpensive, fixed and portable devices [2–4]. These devices can form a sensor network or a Wireless Personal Area Network (WPAN). This last type of network is used for communication among devices (including telephones and personal digital assistants) close to one person. The standard defines a physical layer and a MAC sub layer. Three different frequency ranges are offered. The most important one is the 2.4 GHz range. This is the same range as 802.11b/g and Bluetooth. Consequently, the issue of interference and thus coexistence between the different wireless technologies will be a significant one, especially as reliability is an important requirement. The main contribution of this paper is that we analyze the throughput and delay of IEEE 802.15.4, both analytically and experimentally, for various scenarios such as different addresses and frequency bands. The exact formula for direct communication is drawn up. This gives an overview and an easy way to calculate the maximum throughput without the need to completely analyze the standard. All the information needed for obtaining these results can be found in the standard [5]. The paper is an extension of [6] where only the 2.4 GHz range was considered. In this paper, we will also look into the other frequency bands and offer a more thorough analysis, including the influence of the back off window. Section A.2 of this paper offers a technical overview of the IEEE 802.15.4

frequency band	Symbol rate (baud/s)	Modulation	Bit rate (kbps)
868.0–868.6 MHZ	20 000	BPSK	20
902–928.0 MHz	40 000	BPSK	40
2.4–2.4835 GHz	62 500	16-ary orth.	250

Table A.1: Modulation parameters of IEEE 802.15.4

standard. An overview of related work, such as other performance studies and interference issues, is given in section A.3. In section A.4, the exact formula for the maximum throughput and minimum delay is presented. The analysis of the results is given in section A.5 and experimental validation is done in section A.6. Finally, section A.7 concludes the paper.

A.2 Description of IEEE 802.15.4

The IEEE 802.15.4-standard both defines the physical layer and the medium access layer. For the physical layer, 27 communication channels in three different frequency ranges are defined in the industrial scientific medical (ISM) band: 16 channels in the 2.4 GHz band, 10 channels at 915 MHz and 1 channel at 868 MHz. The 2.4 GHz band is available worldwide and operates at a raw data rate of 250 kbps. The channel of 868 MHz is specified for operation in Europe with a raw data rate of 20 kbps. For North America the 915 MHz band is used at a raw data rate of 40 kbps. An overview of the modulation parameters is given in Table I. All of these channels use DSSS. The standard further specifies that each device shall be capable of transmitting at least 1 mW (0 dBm), but actual transmit power may be lower or higher. Typical devices are expected to cover a 10-20 m range.

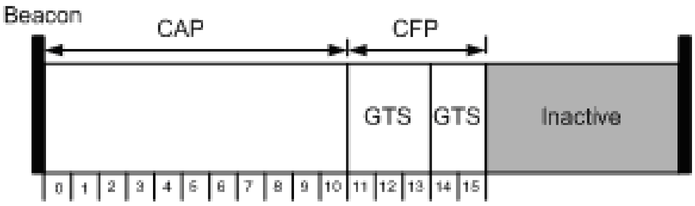


Figure A.1: Superframe structure of IEEE 802.15.4 [5]

An IEEE 802.15.4 network operates either in a beacon enabled mode or in a non beaconed mode. In the beaconless mode, a simple CSMA/CA protocol is used. When a device wishes to transmit data, the device waits for a random number of back off periods. Subsequently, it checks if the medium is idle. If so, the data is

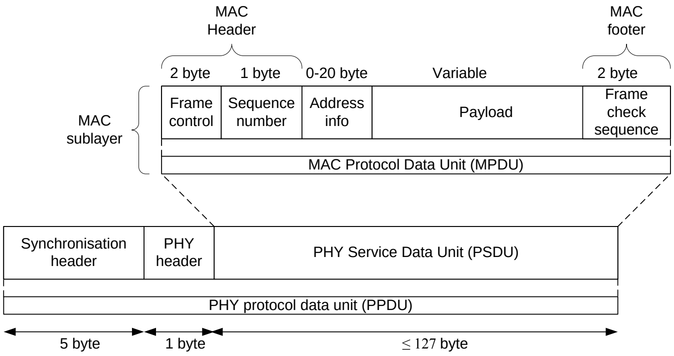


Figure A.2: Frame structure of IEEE 802.15.4

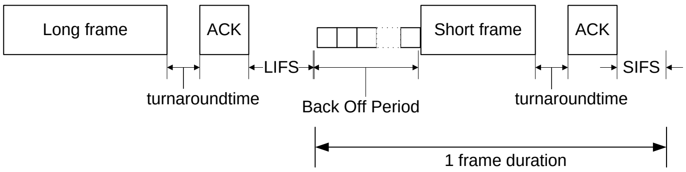


Figure A.3: Frame sequence of IEEE 802.15.4. We notice the back off period and that long frames are followed by a long inter frame space and short frames by a short inter frame space.

transmitted, if not, the device backs off once again and so on. The beaconed mode uses a superframe structure, see Fig. A.1. A superframe starts with beacons sent by a dedicated device, called a coordinator, at predetermined intervals ranging from 15 ms to 251 s. The time between these beacons is split in an active and inactive period. In the inactive period, the device enters a low power mode during which no radio traffic is allowed. Communication between the devices in a PAN (Personal Area Network) only takes place in the active period. The active part is divided in 16 slots of equal size and consists of two groups: the contention access period (CAP) and an optional contention free period (CFP) in order to provide the data with quality of service (QoS). The time slots in the CFP are called guaranteed time slots (GTS) and are assigned by the PAN-coordinator. The channel access in the CAP is contention based (CSMA/CA). By changing the duration of the active and inactive period, the PAN can operate under low duty cycle to save energy.

The MAC sub layer supports multiple network topologies: a star topology with a central network coordinator, a peer to peer topology (i.e. a tree topology) and a combined topology with interconnected stars (clustered stars). The transmitted frames are followed by an Inter Frame Space (IFS) in order to allow the MAC layer a finite amount of time to process data received from the PHY. Before starting the back off period, the device will wait one IFS. Long frames are followed by a Long IFS (LIFS) and short frames by a Short IFS (SIFS). An example of a frame sequence, using acknowledgments (ACKs), is given in Fig. A.2. If no ACKS are used, the IFS follows the frame immediately.

The MAC layer defines either a 64-bit IEEE address or a short 16-bit address assigned during the association period. The packet structure of IEEE 802.15.4 is shown in Fig. A.3. The size of the address info can vary between 0 and 20 bytes as both short and long addresses can be used and as a return acknowledgment frame does not contain any address information at all. Additionally, the address info field can contain the 16-bit PAN identifier, both from the sender and from the receiver. These identifiers can only be omitted when no addresses are sent. The payload of the MAC Protocol Data Unit is variable with the limitation that a complete MAC-frame (MPDU or PSDU) may not exceed 127 bytes. A more thorough description of IEEE 802.15.4 can be found in [5].

A.3 Related Work

Several papers have addressed the issue of performance analysis in IEEE 802.15.4. These papers mainly focus on the slotted version of IEEE 802.15.4 in a multihop environment or with multiple senders and receivers. The performance evaluation of [7] studies the throughput-energy-delay tradeoffs based on NS-2 simulations. It was found that in low duty networks a significant energy saving can be achieved by using the superframe structure, but these savings come at the cost of significantly higher latency and lower bandwidth. A more complete simulation based performance study was done in [8]. An interesting result is that in a non-beaconed mode and for low rate applications the packet delivery ratio of IEEE 802.15.4 is similar to IEEE 802.11. In [9] it was shown that the optimal network performance for slotted CSMA/CA is reached with an offered load in the range of 35% to 60%.

A more theoretical approach was used in [10] where the network is modeled using discrete Markov chains. The throughput and energy consumption were analyzed in saturation conditions for a varying number of nodes in a star topology. In [11] a similar model is defined that considers more parameters. The results show that the average access delays may be quite high if the throughput exceeds 50%.

This paper will focus on the unslotted version of 802.15.4. The approach is similar as the one used in [12] and [13] for IEEE 802.11. A general formula is drawn up and analyzed.

The frequency band with the most number of channels and highest data rates in IEEE 802.15.4 is the 2.4 GHz-band. This is the same band used by IEEE 802.11 (WiFi) [14] and IEEE 802.15.1 (Bluetooth) [15]. These technologies will cause interference when used simultaneously. The interference between WiFi and 802.15.4 was investigated in [16] and [17]. It was concluded that WiFi interference is detrimental to a WPAN using 802.15.4. However, if the distance between the IEEE 802.15.4 and IEEE 802.11b radio exceeds 8 meter, the interference of IEEE 802.11b is almost negligible.

A.4 Calculations

A.4.1 Assumptions

In this paper, we are interested in the throughput of the MAC-layer as seen in the OSI-protocol stack. Therefore, we define the maximum throughput of IEEE 802.15.4 as the number of data bits coming from the upper layer (i.e. the network layer) that can be transmitted. In these theoretical calculations, we will only examine the unbeaconed version of the protocol (i.e. without the superframes). This version has the least overhead so it will give us an upper bound on the maximum throughput of the protocol. The formula will be valid for the different frequency bands. However, the parameters used will have different values at the different frequencies, see section A.5. The maximum throughput is calculated between only one sender and only one receiver which are located close to each other. Hence, we assume that there are no losses due to collisions, no packets are lost due to buffer overflow at either sender or receiver, the sending node has always sufficient packets to send and the BER is zero (i.e. we assume a perfect channel).

A.4.2 Calculations

The maximum throughput (*TP*) is calculated as follows. First the delay of a packet is determined. This overall delay accounts on the one hand for the delay of the data being sent and on the other hand for the delay caused by all the elements of the frame sequence, as is depicted in figure A.3, i.e. back off scheme, sending of an acknowledgement, … In other words, the overall delay is the time needed to transmit 1 packet. Subsequently, this overall delay is used to determine the throughput:

$$TP = \frac{8 \cdot x}{delay(x)} \tag{A.1}$$

In this formula, x represents the number of bytes that has been received from the upper layer, i.e. the payload bytes from figure A.3. The delay each packet experiences can be formulated as:

$$delay(x) = T_{BO} + T_{frame}(x) + T_{TA} + T_{ACK} + T_{IFS}(x) \tag{A.2}$$

The following notations were used:

T_{BO}	=	Back off period
$T_{frame}(x)$	=	Transmission time for a payload of x byte
T_{TA}	=	Turn around time (192 μs)
T_{ACK}	=	Transmission time for an ACK
T_{IFS}	=	IFS time

For the IFS, SIFS is used when the MPDU is smaller than or equal to 18 bytes. Otherwise, LIFS is used. (SIFS = 192 μs, LIFS = 640 μs). The turn around time is the time needed to switch the radio from sending to receiving mode. The different times are expressed as follows:

Back off period:

$$T_{BO} \; = \; BO_{slots} \; \cdot \; T_{BO\,slot} \tag{A.3}$$

BO_{slots}	=	Number of back off slots
$T_{BO\,slot}$	=	Time for a back off slot (320 μs)

The number of back off slots is a random number uniformly in the interval (0, 2^{BE}-1) with BE the *back off exponent* which has a minimum of 3. As we only assume one sender and a BER of zero, the BE will not change. Hence, the number of back off slots can be represented as the mean of the interval: $\frac{2^3-1}{2}$ or 3.5.

Transmission time of a frame with a payload of x bytes:

$$T_{frame}(x) \; = 8 \; \cdot \frac{L_{PHY} + L_{MAC\_HDR} + L_{address} + x + L_{MAC\_FTR}}{R_{data}} \tag{A.4}$$

L_{PHY}	=	Length of the PHY and synch header in bytes (6)
L_{MAC_HDR}	=	Length of the MAC header in bytes (3)
$L_{address}$	=	Length of the MAC address info field
L_{MAC_FTR}	=	Length of the MAC footer in bytes (2)
R_{data}	=	Raw data rate (250 kbps)

$L_{address}$ incorporates the total length of the MAC address info field, thus including the PAN-identifier for both the sender as the destination if addresses are used. The length of one PAN-identifier is 2 bytes.

Transmission time for an acknowledgement:

$$T_{ACK} \; = \; \frac{L_{PHY} + L_{MAC\_HDR} + L_{MAC\_FTR}}{R_{data}} \tag{A.5}$$

If no acknowledgements are used, T_{TA} and T_{ACK} are omitted in (A.2).

Summarizing, we can express the throughput using the following formula:

address bits		868 MHz		915 MHz		2.4 GHz	
		a	b	a	b	a	b
0 bits	ACK	0.0004	0.0149	0.0002	0.00745	0.000032	0.002656
	no ACK	0.0004	0.0099	0.0002	0.00495	0.000032	0.002112
16 bits	ACK	0.0004	0.0181	0.0002	0.00905	0.000032	0.002912
	no ACK	0.0004	0.0131	0.0002	0.00655	0.000032	0.002368
64 bits	ACK	0.0004	0.0229	0.0002	0.01145	0.000032	0.003296
	no ACK	0.0004	0.0179	0.0002	0.00898	0.000032	0.002752

Table A.2: Values for parameters a and b in equations (A.6) and (A.7)

$$TP = \frac{8 \cdot x}{a \cdot x + b} \tag{A.6}$$

$$delay = a \cdot x + b \tag{A.7}$$

In these equations, a and b depend on the length of the data bytes (SIFS or LIFS) and the length of the address used (64 bit, 16 bit or no addresses). The parameter a expresses the delay needed for sending 1 data byte, parameter b is the time needed for the protocol overhead for sending one packet.

A.5 Analysis

In this section, we analyze the throughput, bandwidth efficiency and delay of IEEE 802.15.4 for a number of different scenarios based on the formula above. These scenarios comprehend the use of the different address lengths (64 bit or 16 bit addresses), the use of acknowledgements (ACKs) or not and the different frequencies.

The following parameters from section A.4 are different in the distinct frequency bands:

$$T_{BO\,slot} = 20 \cdot T_S \tag{A.8}$$

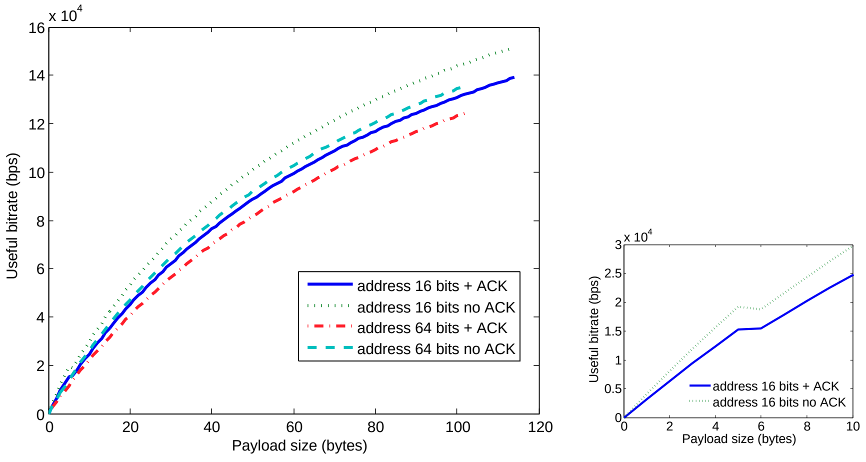
$$T_{TA} = 12 \cdot T_S \tag{A.9}$$

$$T_{SIFS} = 12 \cdot T_S \tag{A.10}$$

$$T_{LIFS} = 40 \cdot T_S \tag{A.11}$$

where T_S represents the duration of one symbol. The value of T_S can be derived from Table A.1 for the different frequency bands.

In Table A.2, the values for a and b are given for the different scenarios, under the assumption that the total packet size is larger than 18 bytes (i.e. LIFS is used). It can be seen that the value of a only depends on the frequency band, which is consistent with the definition of a.



(a) **(b)**

Figure A.4: Useful bitrate in function of a varying payload size for the short and long address scheme, with and without ACK. The frequency is set to 2.4 GHz. Graph (b) shows a snapshot of the left graph for an address size of 16 bits. The transition from SIFS to LIFS can be seen clearly.

A.5.1 Bandwidth Efficiency

The bandwidth efficiency is expressed as

$$BWE = \frac{TP}{R_{data}}. \tag{A.12}$$

We will begin our analysis with the results for the 2.4 GHz band. Fig. A.4 gives the number of useful bits and Fig. A.5 shows the bandwidth efficiency. In the figures, the payload size represents the number of bits that is received from the upper layer. In section II it was mentioned that the maximum size of the MPDU is 127 bytes. Consequently, the number of data bytes that can be sent in one packet is limited. This can be seen in the figures: when the address length is set to 2 bytes (i.e. 16 bits), the maximum payload size is 114 bytes. This can be calculated as follows: MPDU = L_{MAC_HDR} + $L_{address}$ + L_{MAC_FTR}, where $L_{address}$ equals to 2 · 2 bytes + 2 · 2 bytes for the PAN-identifiers and the short addresses respectively. Putting the correct values into the formula for MPDU, gives us 114 bytes. When the long address structure is used (64 bits), 102 data bytes can be put into 1 packet. If no addresses are used, the PAN-identifiers can be omitted, which means that $L_{address}$ is zero. The maximum payload is now set to 122 bytes.

In general, we see that the number of useful bits or the bandwidth efficiency

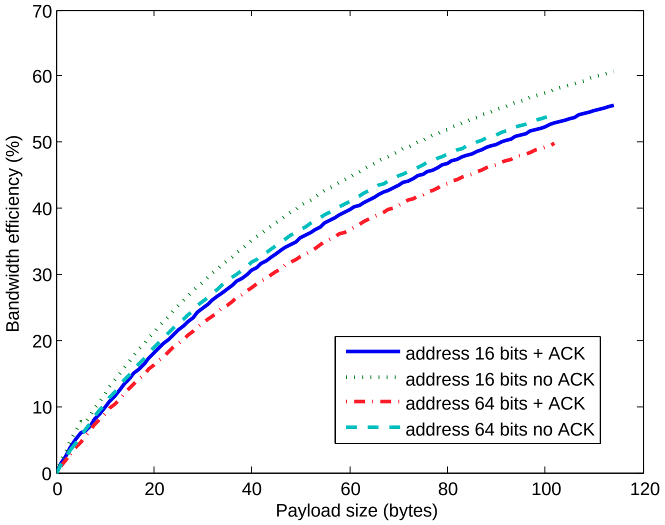


Figure A.5: Bandwidth efficiency for a varying payload size. The frequency is set to 2.4 GHz.

address bits		868 MHz		915 MHz		2.4 GHz	
		Max. rate (bps)	Max. eff.(%)	Max. rate (bps)	Max. effi. (%)	Max. rate (bps)	Max. eff. (%)
0 bits	ACK	16 322	76.6	30 644	76.6	148 780	59.5
	no ACK	16 627	83.1	33 254	83.1	162 234	64.9
16 bits	ACK	14 317	71.6	28 634	71.6	139 024	55.6
	no ACK	15 537	77.7	31 073	77.7	151 596	60.6
64 bits	ACK	12 810	64.1	25 620	64.1	124 390	49.78
	no ACK	13 901	69.5	27 802	69.5	135 638	54.3

Table A.3: Maximum bitrate and bandwidth efficiency of IEEE 802.15.4

grows when the number of payload bits increases. The same remark was made when investigating the throughput of IEEE 802.11 [12, 13] and is to be expected as all the packets have the same overhead irrespective of the length of the packet. Further, the small bump in the graph when the address length is 16 bits at 6 bytes, Fig. A.4(b), is caused by the transition of the use of SIFS to LIFS: at that precise moment the MPDU will be larger than 18 bytes. In all cases, the bandwidth efficiency increases when no ACK is used, which is to be expected as less control traffic is being sent. In Fig. A.4 and A.5 we have only shown the graphs for short and long addresses. The graphs for the scenario without addresses are similar to the previous ones with the understanding that the maximum throughput is higher when no addresses are used. They were omitted for reasons of clarity.

Looking at the figures of Table A.3, we can see that for 2.4 GHz an efficiency of 64.9% can be reached under optimal circumstances, i.e. when no addresses and no acknowledgements are used. If acknowledgements are used, an efficiency of merely 59.5% is obtained. Using the short address further lowers the maximum bit rate by about 4%. The worst result is an efficiency of only 49.8% which is reached when the long address is used with acknowledgements. The main reason for these low results is that the length of the MPDU is limited to 127 bytes. Indeed, the number of overhead bytes is relatively large compared to the number of useful bits (MPDU payload). This short packet length was chosen in order to limit the number of collisions (small packets are used) and to improve fair use of the medium. Further, the main application area of this standard focuses on the transmission of small quantities of data, hence the small data packets.

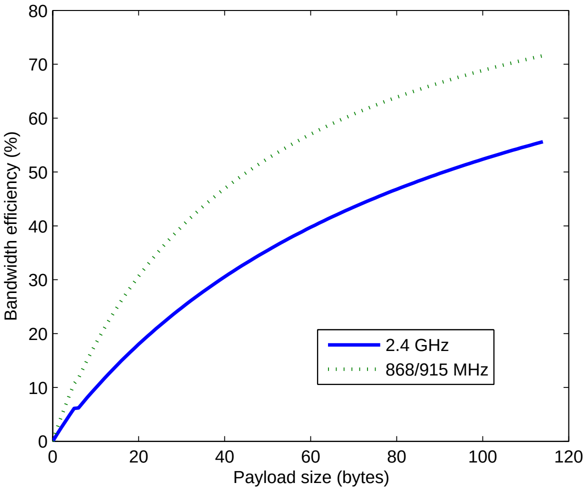


Figure A.6: Bandwidth efficiency for a varying payload size. The frequency is set to 2.4 GHz.

In the other frequency bands, a similar conclusion can be made. Fig.A.6 gives a comparison of the bandwidth efficiency for 2.4 GHz and 868/915 MHz band. A summary can be found in Table A.3 where the maximum bit rate and bandwidth efficiency of the scenarios are given. We see that the bandwidth efficiency is higher for the lower frequency bands and the same for both lower bands. This can be explained as follows. Both bands use BPSK as modulation scheme, see Table I. This means that there is 1 bit per symbol. The 2.4 GHz band on the other hand uses a 16-ary orthogonal modulation with 4 bits per symbol. Some of the time parameters, such as the duration of a SIFS, depend on the duration of 1 symbol, see (A.10). Thus, as the lower bands only have 1 bit per symbol and the 2.4 GHz band 4 bits per symbol, the proportion of the fixed duration to the amount of information that can be sent is higher in the 2.4 GHz band. Consequently, the bandwidth efficiency will be lower in the highest frequency band.

In the lower frequency bands, an efficiency of 83.1% can be reached when no addresses and no ACK are used. The worst result reaches an efficiency of 64.5%. The data rate obtained in the 868 MHz band is exactly half as much as the one in the 915 MHz. Indeed, the baud rate in the 915 MHz band is twice as high as the one in the 868 MHz band (40 kbaud and 20 kbaud respectively) and in both bands 1 bit per symbol is used.

A.5.2 Delay

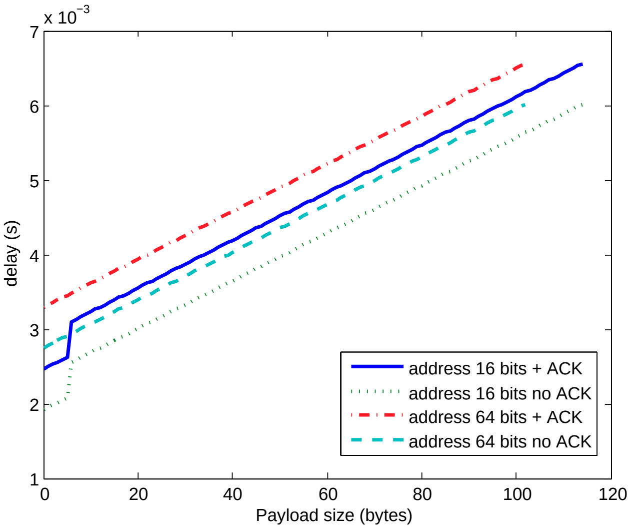


Figure A.7: Minimum delay as a function of the payload size. The frequency is set to 2.4 GHz.

Fig. A.7 gives the minimum delay each packet experiences for varying packet sizes in the 2.4 GHz band. Table A.4 gives the minimum and maximum delay for the different scenarios. The minimum delay is calculated by sending a packet without any data bits immediately from 1 sender to 1 receiver. In other words, the minimum delay is the time needed to send and process an empty packet in a single hop environment. The propagation delay is not taken into account and no retransmissions are assumed. We immediately notice that the delay is a linear function of the number of payload bytes, as long as we assume a payload of more than 10 bytes. The jump in the graph for the short address length is caused by the IFS-mechanism. The same behavior is found for the other frequency band.

address bits		868 MHz		915 MHz		2.4 GHz	
		Min. delay	Max. delay	Min. delay	Max. delay	Min. delay	Max. delay
0 bits	ACK	13.5	63.7	6.75	31.85	2.21	6.56
	no ACK	8.5	58.7	4.25	29.35	1.66	6.02
16 bits	ACK	16.7	63.7	8.35	31.85	2.46	6.56
	no ACK	11.7	58.7	5.85	29.35	1.92	6.02
64 bits	ACK	22.9	63.7	11.45	31.85	3.30	6.56
	no ACK	17.9	58.7	8.95	29.35	2.75	6.02

Table A.4: Minimum and maximum delay of IEEE 802.15.4 (ms)

The maximum delay is found by sending a full packet, i.e. the MPDU is set to the maximum of 127 bytes. This means that the maximum payload is independent of the number of address bits. However, as can be seen in Fig. A.7, the maximum number of payload bits differs when the short or long address is used. We see that the maximum delay is a little bit higher than 6 ms in the 2.4 GHz region when a full packet is sent. This delay is acceptable for delay bound applications. The lower bands experience a significant higher delay, which is to be expected as the data rate is lower. In these frequency bands it is more important to look to the minimum delay, especially in the 868 MHz band. These figures can offer a more thorough insight in the limitations of IEEE 802.15.4 when designing and implementing a network based on this protocol. For example, if one plans to use the 868 MHz band in order to lower the coexistence issues, it is important to know the minimum delay bounds.

A.5.3 Influence of Back Off Window

The IEEE 802.15.4 protocol uses CSMA/CA. This means that when a device wants to transmit data, the device waits for a random number of back off periods before trying to access the channel. The back off time is randomly generated in the interval [0 , 2^{BE}-1], see section IV. The maximum value of BE is 5, which means that the number of back off slots is limited to 31. This is significantly lower than the maximum number of 1023 back off slots in IEEE 802.11 [14]. This will influence the network throughput when multiple radios are used as collisions will be more likely due to the small back off interval. The initial value of BE at the first back off is called macMinBE and has a default value of 3. However, this value can be altered in the range of [0,5]. When macMinBE is set to 0, no collision avoidance is done in the first attempt to access the channel.

The macMinBE value has a significant influence on the bandwidth efficiency. Fig. A.8 shows the bandwidth efficiency for the frequency bands of 2.4 GHz and 915 MHz. 16 bit addresses are used without acknowledgement. When the

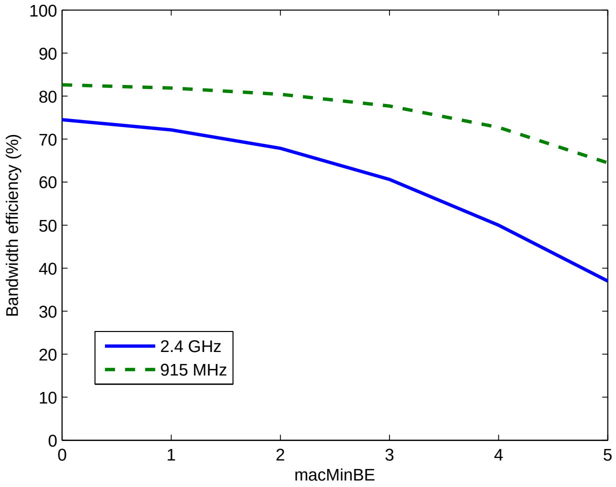


Figure A.8: The bandwidth efficiency for different values of *macminBE* in the 2.4 GHz and 915 MHz band with short addresses and without acknowledgments.

macMinBE is set to 0, an efficiency of 74.5% is reached (or 186 270 bps), when set to 5, the efficiency is only 37.0% (or 92 532 bps). The default case resulted in an efficiency of 60.6%, see Table II. The difference is quite large. If we look at the scenario where no addresses are used and without acknowledgement, we see an efficiency of 79.7% when no collision avoidance is used. In the default case, only 64.9% was reached, Table II. Therefore, it is interesting to lower the macMinBE value when only one sender and receiver are used. Further, it can be seen that the 2.4 GHz suffers more from an increasing value of macMinBE. This behavior is expected as the 2.4 GHz band uses more bits per symbol.

Now let's consider a scenario with multiple senders and receivers. In dense networks with high traffic loads, a lot of collisions will occur and the nodes will back off several times. The maximum back off time is 31 and is reached after 2 back offs if the default value is used and after 5 back offs when no collision avoidance is done on the first attempt. Consequently, most of the nodes will have the maximum back off interval and the influence of differentiating the macMinBE will be limited. The (aggregated) throughput of the nodes will decrease due to the large back off interval. In low density networks, the probability of a collision will be significantly lower. Therefore it would be interesting to lower the macMinBE in such scenarios.

A.6 Experimental Results

In our analysis above, we have determined the theoretical throughput of IEEE 802.15.4. In this section, we compare this analysis with experimental results. We have measured the throughput between two radios that use the IEEE 802.15.4 specification.

For our assays, we have used the 13192 DSK (Developer's Starter Kit) of Freescale Inc. This kit uses the MC13192 radio chip of Freescale Inc. [18]. The radio operates at 2.4 GHz and libraries are included which implement the IEEE 802.15.4-standard. We used the highest channel (channel 26) as this channel does not overlap with any of the channels of IEEE 802.11 [14]. This way we minimize the interference caused by the 802.11 radio transmission. This channel does still overlap with the Bluetooth spectrum [15]. However, Bluetooth uses a frequency hopping technique, so the interference will be limited to a short time. Therefore we have measured over a period of 1 day. Although we have tried to minimize the interference, not all of the packets were received correctly. Consequently, some of the packets needed to be retransmitted. This not only causes an extra delay due to the second transmission, but also an increase of the size of the back off window. This mechanism will negatively influence the experimental throughput. We have placed the sender and receiver 1 meter apart at a height of 1 meter.

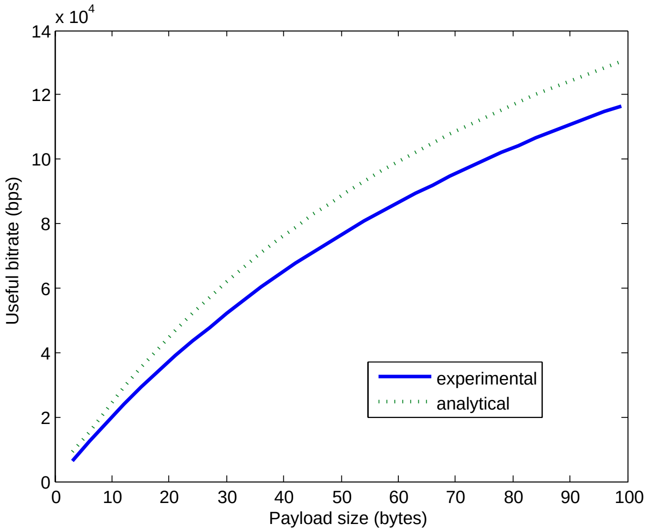


Figure A.9: Comparison between analytic result and experiment.

Fig. A.9 gives a comparison between the theoretically and experimentally obtained results when a short address and an acknowledgment are used. Once again, we varied the number of payload bytes. We see that the experimental curve is lower than the one obtained analytically. This is to be expected as the theoretical analysis offers an upper bound to the throughput. We further notice that the 2 graphs have the same curve. The relative difference between the two curves is steady at about 11%. We have fitted the experimental curve with (A.1) and obtained the following values for a and b respectively: 0.0000324 and 0.00359. The theoretical values can be found in table II (16 bit address and ACK used). It can be seen that the main difference is to be found in the part that is independent of the number of bytes sent. This is an indication that an extra delay or processing time needs to be added to each packet. The duration of this extra delay is about 680 μs (b is expressed in seconds: 0.00359-0.00291 = 0.00068 s).

We also measured the throughput for the scenario where long addresses are used without acknowledgments. Again a lower throughput than theoretically expected is achieved. Now we see a difference of about 9%. The fitted values for a and b are 0.00003201 and 0.003721 respectively. The extra delay is about 520 μs and once again independent of the number of bits sent. The time difference between the two scenarios is comparable. The experimental results match better when no ACKs are used. This is to be expected as no retransmissions will occur when packets get lost.

This allows us to conclude that the difference between the experimental and theoretical results is mainly due to delays caused by the electronics of the radios used and the occurrence of retransmissions.

A.7 Conclusion

The maximum throughput and minimum delay were determined under the condition that there is only one radio sending and one radio receiving. The next step in analyzing the performance of IEEE 802.15.4 would be introducing more transmitters and receivers which can hear each other. It is assumed that the maximum overall throughput, i.e. the throughput of all the radios achieved together, will fall as the different radios have to access the same medium which will result in collisions and longer back off periods. This also will result in lower throughput and larger delays.

We have presented the exact formula for determining the maximum throughput of the unbeaconed version of IEEE 802.15.4 for different frequency bands and scenarios. It was concluded that the throughput varies with the number of data bits in the packet. In the 2.4 GHz band a maximum throughput of 163 kbps or an efficiency of 64.9% can be achieved. The other frequency bands offer a higher efficiency, but a lower effective throughput. By changing the back off exponent, a higher throughput can be obtained. It is concluded that the bandwidth efficiency is rather low due to the small packet size imposed in the standard.

References

[1] ZigBee Alliance, official webpage: http://www.zigbee.org.

[2] E. Callaway, P. Gorday, L. Hester, J. A. Gutierrez, M. Naeve, B. Heile, and V. Bahl. *Home networking with IEEE 802.15.4: a developing standard for low-rate wireless personal area networks*. IEEE Communications Magazine, 40(8):70–77, August 2002.

[3] J. Zheng and M. J. Lee. *Will IEEE 802.15.4 make ubiquitous networking a reality?: a discussion on a potential low power, low bit rate standard*. IEEE Communications Magazine, 42(6):140–146, June 2004.

[4] J. A. Gutierrez, M. Naeve, E. Callaway, M. Bourgeois, V. Mitter, and B. Heile. *IEEE 802.15.4: a developing standard for low-power low-costwireless personal area networks*. IEEE Network, 15(5):12–19, September/October 2001.

[5] IEEE 802.15.4-2003: IEEE Standard for Information Technology - Part 15.4: Wireless Medium Access Control and Physical Layer specifications for Low Rate Wireless Personal Area Networks.

[6] B. Latré, P. De Mil, I. Moerman, N. Van Dierdonck, B. Dhoedt, and P. Demeester. *Maximum Throughput and Minimum Delay in IEEE 802.15.4*. Lecture Notes in Computer Science , Proceedings of the 1st International Conference on Mobile Ad-hoc and Sensor Networks, 3794:866–876, December 2005. ISSN 1796-2056.

[7] G Lu, B Krishnamachari, and C S. Raghavendra. *Performance Evaluation of the IEEE 802.15.4 MAC for Low-Rate Low-Power Wireless Networks*. In Workshop on Energy Efficient Wireless Communications and Networks (EWCN'04), April 2004.

[8] J. Zheng and Myung J. Lee. *Sensor Network Operations*, chapter A comprehensive performance study of IEEE 802.15.4, pages 218–237. IEEE Press, Wiley Interscience, 2006.

[9] A. Koubaa, M. Alves, and E. Tovar. *A comprehensive simulation study of slotted CSMA/CA for IEEE 802.15.4 wireless sensor networks*. In Factory Communication Systems, 2006 IEEE International Workshop on, pages 183–192, June 2006.

[10] T. R. Park, T. H. Kim, J. Y. Choi, S. Choi, and W. H. Kwon. *Throughput and energy consumption analysis of IEEE 802.15.4 slotted CSMA/CA*. Electronics Letters, 41(18):1017–1019, September 2005.

[11] J. Mišic and V. B. Mišic. *Access Delay and Throughput for Uplink Transmissions in IEEE 802.15.4 PAN*. Elsevier Computer Communications Journal, 28(10):1152–1166, June 2005.

[12] Y. Xiao and J. Rosdahl. *Throughput and delay limits of IEEE 802.11*. IEEE Communications Letters, 6(8):355–357, August 2002.

[13] J. Jun, P. Peddabachagari, and M. Sichitiu. *Theoretical maximum throughput of IEEE 802.11 and its applications*. In Network Computing and Applications, 2003. NCA 2003. Second IEEE International Symposium on, pages 249–256, April 2003.

[14] IEEE 802.11-1999: IEEE Standard for Information Technology - Part 11: Wireless LAN Medium Access Control and Physical Layer specifications.

[15] P. Johansson, M. Kazantzidis, R. Kapoor, and M. Gerla. *Bluetooth: an enabler for personal area networking*. IEEE Network, 15(5):28–37, September/October 2001.

[16] N. Golmie, D. Cypher, and O. Rebala. *Performance analysis of low rate wireless technologies for medical applications*. Computer Communications, 28(10):1266–1275, June 2005.

[17] Dae Gil Yoon, Soo Young Shin, Wook Hyun Kwon, and Hong Seong Park. *Packet Error Rate Analysis of IEEE 802.11b under IEEE 802.15.4 Interference*. In Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd, volume 3, pages 1186–1190, Melbourne, Vic.,, May 2006.

[18] Freescale Inc. [Online] http://www.freescale.com/zigbee.

B

The Need for Cooperation and Relaying in Short-Range High Path Loss Sensor Networks

B. Braem[1], B. Latré[2], I. Moerman[2], C. Blondia[1], E. Reusens[3], W. Joseph[3] and P. Demeester[2]

published in Proceedings of the IEEE International Conference on Sensor Technologies and Applications (Sensorcomm 2007), Valencia, Spain, October 2007, pp. 566–571

[1] University of Antwerp – IBBT, Dept. of Mathematics and Computer Science, PATS research group, Antwerp, Belgium
[2] Ghent University – IBBT, Dept. of Information Technology, IBCN research group, Gent, Belgium
[3] Ghent University – IBBT, Dept. of Information Technology, WiCa research group, Gent, Belgium

Abstract *This paper focuses on the energy efficiency of communication in small-scale sensor networks experiencing high path loss. In particular, a sensor network on the human body or BASN is considered. The energy consumption or network lifetime of a single-hop network and a multi-hop network are compared. We derive a propagation model and a radio model for communication along the human body. Using these models, energy efficiency was studied analytically for* *a line and a tree topology. Calculations show that single-hop communication is inefficient, especially for nodes far away from the sink. There however, multi-hop proves to be more efficient but closer to the sink hotspots arise. Based on these findings, we propose to exploit the performance difference by either introducing extra nodes in the network, i.e. dedicated relay devices, or by using a cooperative approach or by a combination of both. We show that these solutions increase the network lifetime significantly.*

B.1 Introduction

Sensor networks are an interesting application of recent wireless technology. The classical scope of these networks are large-field setups where lots of nodes are scattered around the area being monitored. In this paper we consider the entirely different area of small-scale and short-range sensor networks. In particular, we will look into body area sensor networks or BASNs [1, 2]. In these networks, sensors are attached to the human body, they collect information about the person and send it wirelessly to the sink, a device that acts as a gateway to other networks or processes the data. BASNs can be large, as they should support athletes with lots of sensors attached to their body e.g. movement sensors on limbs.

Energy consumption is a large issue in BASNs, as it is in regular sensor networks. It is not possible to equip the sensors with replaceable or rechargeable batteries as this reduces the comfort of the person wearing them [3]. Further, in a BASN communication takes place near the human body which is a very lossy medium. Consequently, the electromagnetic waves are attenuated considerably, or stated otherwise, the radio signals experience a high path loss. This means that transmitting over an arbitrary distance near the human body is not always possible. Another problem may be possible tissue heating [4, 5]. This effect can arise when too much power is transmitted near the human body. Regulation similar to the one for mobile phones is in place, with strict transmit power requirements [6, 7]. Combined with the higher path loss, these results motivate the use of multi-hop networks.

In section B.2, an overview of research about the path loss along the human body is given. It is clear that high path losses are experienced. Afterwards an overview of existing radio models used in sensor networks is presented and existing work on the multi-hop communication in BASN is considered. Section B.3 explains the propagation model and the radio model that are used in our analysis. A first comparison between single-hop and multi-hop communication is made in section B.4. Mechanisms to improve the energy efficiency of sensor networks based on the previous results are proposed in section B.5. Finally, section B.6 gives directions for future work and section B.7 concludes this paper.

B.2 Related Work

B.2.1 Path Loss Models for the Human Body

Several researchers have been investigating the path loss along and inside the human body either using narrowband radio signals or Ultra Wide Band (UWB). All of them come to the conclusion that the radio signals experience great losses and that the value of the path loss exponent η varies greatly. The propagation of electromagnetic waves in the human body, where the tissue medium acts as a communication channel, has been investigated in [5, 8]. It is concluded that the path loss is very high compared to free space propagation. The channel model for line of sight (LOS) propagation along the human body was studied in [9, 10]. It was found that the path loss exponent is about 3. In [11], a path loss exponent of 7 was found in non-line of sight (NLOS) situations for propagation along the body. This means that the path loss around the human body may thus tremendously exceed the path loss for propagation in free space ($\eta = 2$). Due to these high losses, direct communication between the sensors and the sink will not always be possible, especially when one wants to lower the radio's transmission power. Hence, multi-hop networking becomes advantageous and sometimes even an absolute requirement to ensure connectivity of the network.

B.2.2 Radio Models

An important element in analyzing the energy efficiency of a network, is to have a good radio model at one's disposal. As we are only interested in the energy consumption of the communication, which is much larger than the energy used for sensing [12], we ignore the latter in this paper. Different radio models can be found in the literature. In [13] a first order radio model is proposed. The model assumes a d^2 energy loss due to channel transmission with d the distance between sender and receiver.

$$E_{tx}(k,d) = E_{TXelec} \cdot k + E_{amp} \cdot k \cdot d^2 \tag{B.1}$$

$$E_{rx}(k) = E_{RXelec} \cdot k \tag{B.2}$$

In these formulas, E_{tx} represents the transmission energy, E_{rx} the receiver energy, E_{TXelec} and E_{RXelec} the energy the radio dissipates to run the circuitry for the transmitter and receiver respectively, E_{amp} the energy for the transmit amplifier and k the number of bits sent. The radios have power control and consume the minimal energy needed to reach the receiver. A drawback of this model is the assumption that the transmitter is able to perform power control, which is not as simple as it seems.

In [14] a model is presented where the node decrements the available energy according to the following parameters: (a) the specific network interface controller characteristics, (b) size of the packets and (c) the bandwidth used. The following equations represent the energy used (in Joules) when a packet is transmitted (B.3) or received (B.4). The packet size is in bits.

$$E_{transmitting} = \frac{I_{transmitting} \cdot V}{Bandwidth} \cdot Packetsize \tag{B.3}$$

$$E_{receiving} = \frac{I_{receiving} \cdot V}{Bandwidth} \cdot Packetsize \tag{B.4}$$

Although the equipment not only consumes energy when sending and receiving but also when listening, the models above assume that the listening operation is energy free. This model does not take a static energy consumption due to processing packets into account.

B.2.3 Multi-hop Communication in BASNs

Most researchers in the area of communication in a BASN only consider single-hop communication between the sensors and the sink. [15, 16] define a relatively simple Time-Division Multiple-Access (TDMA) protocol and an adapted implementation of IEEE 802.15.4 is used in [17]. Very few analysis about multi-hop communication has been done. In [18] a first attempt was made to justify the use of multi-hop networking when using UWB communication. It was concluded that the criteria whether to use a multi-hop strategy depend on the ratio of the energy consumption needed for decoding/coding and receiving/generating a UWB-pulse. A preliminary research for narrowband communication was done in [19]. This work only considered the energy consumption of the entire network, it did not take the individual nodes into account.

B.3 Propagation and Radio Model

In order to evaluate the energy consumption in short range wireless networks, we need to select a propagation model and a radio model. To model the propagation between the transmitting and the receiving antenna as a function of the distance d, we use the following semi-empirical formula for the path loss, presented in [9, 11]:

$$P_{dB} \; = \; P_{0,dB} \; + \; 10 \cdot \eta \cdot \log_{10}(\frac{d}{d_0}) \tag{B.5}$$

where $P_{0,dB}$ is the path loss at a reference distance d_0 and η is the path loss exponent, which equals 2 in free space.

Table B.1 shows the parameter values of the fitted path loss models for two different propagation channels, according to equation (B.5), and the variation σ of the individual measurements around the model. The first channel is located along the front of the torso and is LOS [9]. The second channel is measured around

parameter	value LOS [9]	value NLOS [11]
d_0	10 cm	10 cm
$P_{0,dB}$	35.7 dB	48.8 dB
σ	6.2 dB	5.0 dB
η	3.38	5.9

Table B.1: Parameter values for the path loss model

the torso, resulting in NLOS propagation [11]. We observe a higher path loss and higher path loss exponent along the NLOS channel than along the LOS channel, due to diffraction around the human body and absorption of a large amount of radiation by the body.

The model in equation (B.5) only represents the mean path loss [20]. In practice, there will be variations with respect to the nominal value. This variation is well described by a lognormal distribution, and is called *shadowing*. It is crucial to account for this in order to provide a certain reliability of communications. The total path loss then becomes a random variable given by

$$PL = PL_{dB} + PL_s \tag{B.6}$$

where PL_{dB} is the value predicted by the path loss model (B.5), and the shadowing component PL_s is a zero-mean Gaussian random variable with standard deviation σ (see Table B.1). In order to provide reliable communications, the extra margin $PL_s = t \cdot \sigma$ has to be added, according to the reliability required from the system. The value of t can be calculated according to this formula:

$$t = \sqrt{2} \cdot erfc^{-1}[2 \cdot (1 - p)] \tag{B.7}$$

where $erfc^{-1}()$ is the inverse of the standard cumulative error function, and p is the percentage of reliability that is required. For example, if we want to obtain a reliability of 99%, which seems suitable for reliable body area sensor networks, the value of t is 2.326.

As radio model, we have chosen the first order model described in [13] and section B.2.2. In order to allow for a more general use of the formula, we change it to d^η where η is the loss coefficient. Further, E_{amp} varies according to the loss coefficient, so we used $E_{amp}(\eta)$ instead. The specific values of these parameters are hardware dependent. We have determined these parameters for 2 commercially available transceivers which are frequently used in sensor networks: the Nordic nRF2401 low power single chip transceiver [21] and the Chipcon CC2420 transceiver [22] used in Telos-B motes. Both transceivers work in the 2.4–2.45 GHz band and have a very low power consumption. The appropriate values for the parameters above were obtained by fitting (B.1) and (B.2) to the actual power

parameter	nRF2401	CC2420
E_{TXelec}	16.7 nJ/bit	96.9 nJ/bit
E_{RXelec}	36.1 nJ/bit	172.8 nJ/bit
$E_{amp}(3.38)$	1.97e-9 J/bit	2.71e-7 J/bit
$E_{amp}(5.9)$	7.99e-6 J/bit	9.18e-4 J/bit

Table B.2: Parameter values for the Nordic nRF2401 and Chipcon CC2420 transceivers

consumption of the devices which can be found in the datasheets. The distance used in (B.1) is the maximal distance that can be reached between the sender and the receiver. If the receiver is positioned a little bit further, it can no longer hear the sender. This distance is calculated for each output power level (P_{tx}) using (B.5) and the assumption that the maximal path loss (P_{dB}) equals the difference between the sensitivity of the radio and P_{tx}. The results for both radios for different values of the path loss exponent can be found in Table B.2. It can be seen that the Nordic radio has a lower energy consumption per bit. This can be explained by the higher bitrate that can be obtained by the Nordic transceiver. Hence, we will use the parameters of the Nordic radios in our further calculations.

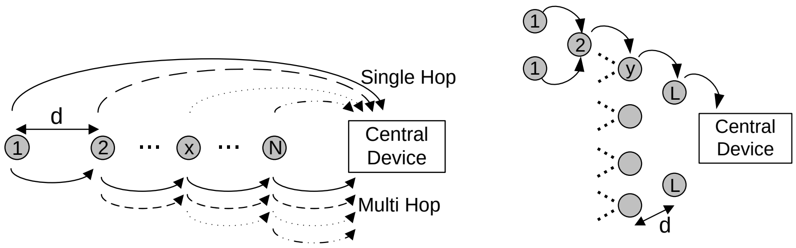
B.4 Single-hop versus Multi-hop

To study the effects of a multi-hop approach we take two different topologies into account: in the first one all the nodes are equidistantly placed on one line (Line topology) and in the second one the nodes form a tree network (Tree topology). In both we assume that all nodes in the network generate packets at the same rate, so each duty cycle each node wants to send one packet to the sink. Based on the radio model and the propagation model described in section B.3, we used a smaller pathloss exponent for links between nearby nodes. Consequently all single-hop transmissions use a high pathloss, i.e. the path loss exponent of NLOS situations, except for the nodes next to the sink which use the LOS value. In the multi-hop scenario, the LOS value is used for transmission to neighboring nodes. In this study, a perfect duty cycle is assumed, i.e. a sensor only turns on its radio when it sends or receives data. The main purpose of this approach is to orthogonalize the results from this study and the properties of specific MAC-protocols.

As the energy efficiency is considered as one of the most important performance issues of BASNs, we use the network lifetime as metric, which we define as the time for the first node to die. In order to have a high network lifetime, the most consuming node should be made more energy efficient. This metric forces us to consider all nodes to be equally important, which corresponds to the fact that the sensors generate and transport medical data.

B.4.1 Line Topology

A first topology we considered is very simple: all nodes are on one line, as shown on figure B.1(a). The distance between the nodes is fixed at d. The counting starts from the node the farthest from the sink.



(a) A line topology

(b) A tree topology

Figure B.1: Examples of topology

When we have N nodes in the network, we can write the energy usage per bit for node x when using single-hop as

$$\begin{aligned} E_{SH}(x,d) = {} & E_{TXelec} \\ & \ldots + \; E_{amp}(\eta) \; \cdot \; ((N - x + 1\,) \cdot d)^{\eta} \end{aligned} \tag{B.8}$$

Whereas the energy usage in a multi-hop network is somewhat more complicated and a term for the energy consumption while receiving is added:

$$\begin{aligned} E_{MH}(x,d) = {} & (x-1) \cdot E_{RXelec} \; + \\ & \ldots x \cdot (E_{TXelec} \; + \; E_{amp}(\eta) \; \cdot \; d^{\eta}) \end{aligned} \tag{B.9}$$

Figure B.2 shows the ratio of single-hop energy usage over multi-hop energy usage for a scenario with 4 nodes and different pathlosses, i.e. we use the LOS and NLOS values. The results show that the nodes closest to the sink perform really bad when using multi-hop: they become *hotspots* using more than 10 times the energy of single-hop because they are relaying a lot. However, far away from the sink, at node 1, single-hop performs up to 1000 times worse because of the high pathloss. It is clear that distance plays an important role in these results. When the distance between the node increases, the single-hop path loss effects start to impact the performance dramatically.

When looking at larger networks of 5 and 6 nodes in figure B.3 we notice the same pattern: the 2 nodes closest to the sink perform really bad in multi-hop, the nodes far away from the sink perform a lot worse in the single-hop approach compared to the multi-hop one.

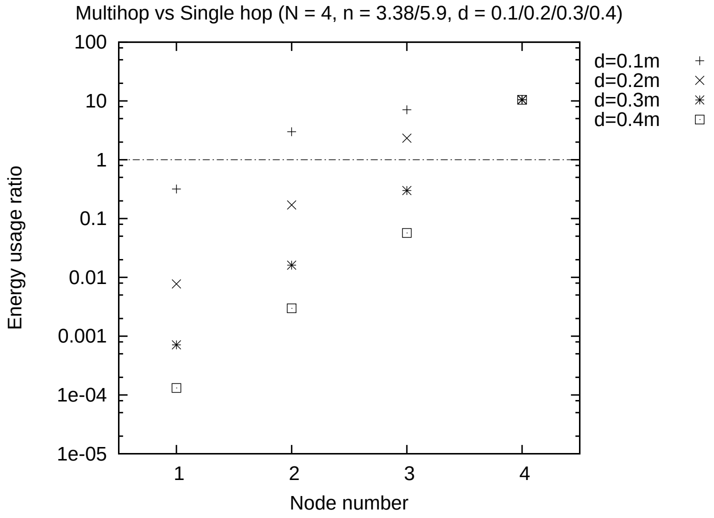


Figure B.2: Energy usage ratio in a line scenario with 4 nodes

B.4.2 Tree Topology

This topology is commonly used in sensornetworks and induces a larger forwarding overhead because of the increased number of children. In BASNs the formation of the different levels largely depends on the situation: the number of nodes in the network, how far the nodes are placed apart,... In this paper, we do not consider how the levels are formed. Only the generic case of a full binary tree is considered. Each node in the tree has exactly one parent and two children, as depicted in figure B.1(b). The choice of a full binary tree is arbitrary, but allows an easier analytical evaluation. However, it should be mentioned that a tree with 7 hops, i.e. 7 levels, consists of 127 nodes. Thus, the network is largely overdimensioned to simulate large loads. When there are L levels in the network, in this topology the energy usage per bit for a node at level y when using single-hop can be written as

$$E_{SH}(y,d) = E_{TXelec} + E_{amp}(\eta)\,((L-y+1)\cdot d)^{\eta} \tag{B.10}$$

Whereas the energy usage for a node at level y in a multi-hop network is given as:

$$\begin{aligned} E_{MH}(y,d) &= (2^y-2)\cdot E_{RXelec} + \\ &\quad \ldots(2^y-1)\cdot(E_{TXelec} + E_{amp}(\eta)\cdot d^{\eta}) \end{aligned} \tag{B.11}$$

When looking at the performance ratios in figure B.4, the situation looks quite similar. The tree topology and the resulting higher forwarding overhead makes the nodes near the sink perform even worse, further away from the sink the single-hop situation does not change.

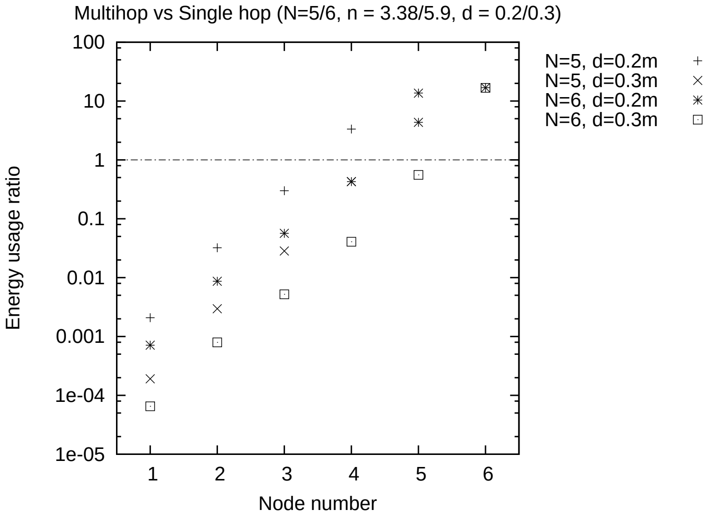


Figure B.3: Energy usage ratio in a line scenario with 5 and 6 nodes

As this case is more general, we will try to improve performance for this scenario. Any improvements will then be trivially the same for the single-hop scenario.

B.5 Improving the Energy Efficiency

The results obtained in the previous section are used to improve the energy efficiency of communication taking place in BASNs. As stated before, we consider the network lifetime to be the most important metric. It can be improved by tackling the energy usage at the nodes consuming the most energy.

If we look at both the line and the tree topology, we see that in single-hop there is clearly room for energy saving at the nodes further away from the sink. These nodes consume the most energy and consequently will die first. However, we also see that in the multi-hop scenario, more energy is consumed by the nodes closest to the sink as they have to forward the data received from nodes farther away. Based on these observations, in this section we will propose 2 mechanisms that can be used in order to improve the network lifetime considerably: relaying and cooperation.

B.5.1 Relaying

A first solution encompasses the introduction of dedicated relay devices. These are special nodes which only handle traffic relaying and do not do any sensing

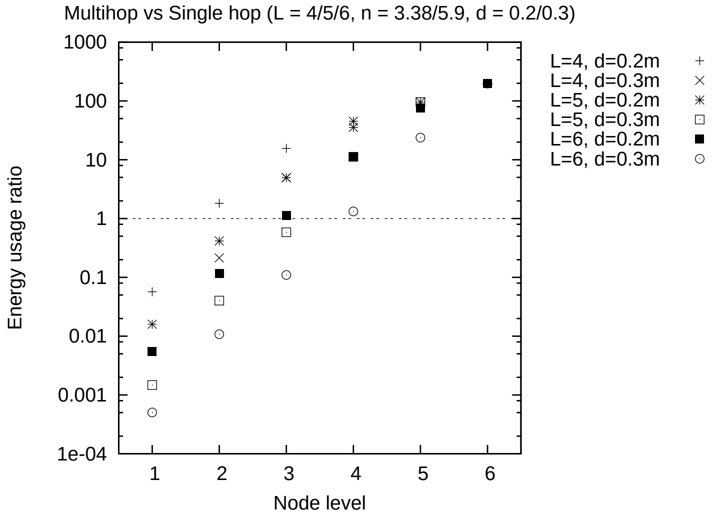


Figure B.4: Energy usage ratio in a tree scenario with 4,5 and 6 levels

themselves, thus more energy is available for communication purposes. The main idea is that proper placement of relay nodes can bridge the performance gap for the nodes far away from the sink in the case of single-hop traffic and offload the nodes closer to the sink in the case of multi-hop traffic.

For a node relaying traffic from z nodes, energy usage per bit is similar to a regular node in a multi-hop tree network (B.11), minus the cost of transmitting own packets:

$$\begin{aligned} E_R(z,d) &= z \cdot (E_{RXelec} + \\ &\quad \ldots (E_{TXelec} + E_{amp}(\eta) \cdot d^{\eta})) \end{aligned} \tag{B.12}$$

In this formula, d represents the distance to the next relay hop.

As an example, we consider a tree network of 5 hops. Relay devices are placed at level 4 and relay directly to the sink. Figure B.5 shows the result when the distance between the nodes is 20 cm and figure B.6 gives the result for a network with 30 cm between the nodes. The graphs for single-hop and multi-hop communication are plotted. Further, the energy consumption of the relayed network is shown: the nodes of level 1 and level 2 all send their data to the relay device at level 4. It can be seen that the lifetime of the nodes at level 1 and 2 improves a lot with respect to the single-hop scenario. In both cases the energy usage at level 1 decreases by a factor 20.

The points at level 4 represent the energy usage of a relay node when it forwards 1, 3 or 7 nodes. When the distance between the nodes is 20 cm (figure B.5), we see that the energy needed for relaying 3 nodes is lower than the energy usage at level 3. If the distance is larger, i.e. 30 cm, it is even possible to relay 7 nodes while staying under the energy consumption at level 3. However, depending on the energy required for sensing, supporting a larger number of nodes should be possible.

It should be noted that the number of nodes in this example network is very high, the number of relay nodes will not have to be as high in realistic networks.

When considering networks with more hops, the introduction of relay devices is clearly better because of the high path loss. The position of the relay nodes is highly situation dependent. Yet, the following rule of thumb can be used: the placement should not be too far away from the sink as the path loss effects will impact efficiency dramatically. A position closer to the sink is a better option, however the number of hops between the nodes and the relay device should not become too large.

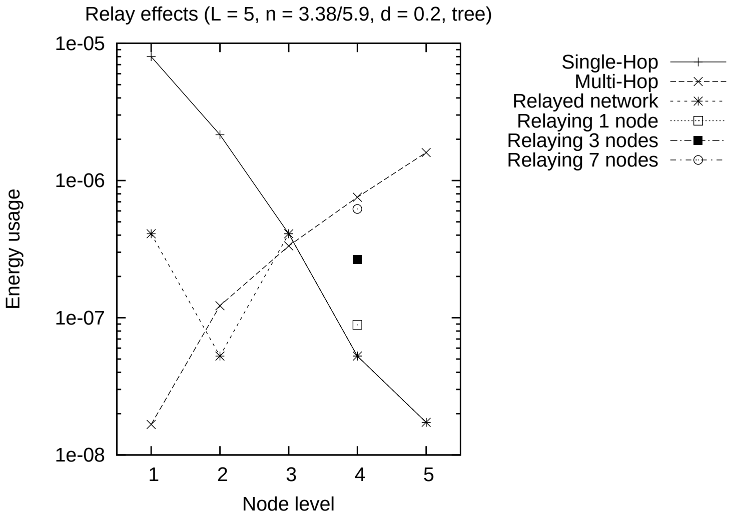


Figure B.5: Energy usage when relaying with inter-node distance 20 cm

B.5.2 Cooperation

The previous section shows that using relay nodes considerably improves the lifetime of the network. However, it is not always feasible to use relay nodes. Specifically in the case of body area sensor networks, putting even more sensors on users does not really improve comfort. Hence, other methods need to be found.

When we look at figures B.5 and B.6, it is obvious that there is a lot of residual energy available at levels 4 and 5 compared with the energy usage of the nodes

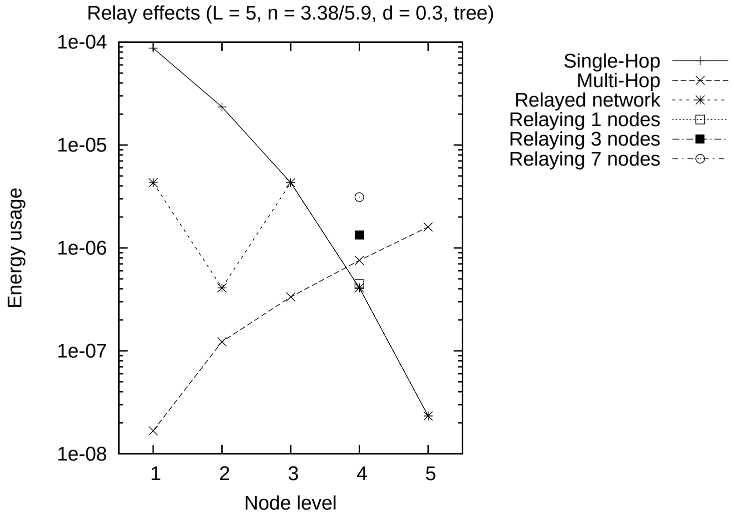


Figure B.6: Energy usage when relaying with inter-node distance 30 cm

at level 3. The solution proposed is to use this residual energy for relaying data from other nodes. Stated otherwise, to let those nodes cooperate in the network. Indeed, by breaking into this energy supply, the lifetime of the network will still be bounded by the energy usage of the nodes on level 3. The data of the nodes on level 1 and 2 can be forwarded to level 4 and level 5 respectively. This will lower the energy consumption of these nodes, as was the case when using relay devices. Hence, the network lifetime can be improved without the addition of extra relay devices.

In the smaller network, i.e. when the nodes are placed 20 cm apart, it can be calculated that the data of up to 4 nodes can be relayed by the nodes at level 4. The following formula is used for level k when the energy consumption is limited by level l:

$$\#\text{nodes supported} = \lfloor \frac{E_{SH}(l,d) - E_{SH}(k,d)}{E_R(1,(L-k+1)\cdot d)} \rfloor \tag{B.13}$$

The energy consumption of a node at level 4 or 5 sending its own data and relaying the data of the other nodes still remains below the energy consumption of node 3. Thus, the lifetime of the network remains the same. Figure B.7 shows an example of the results when using this approach. The almost horizontal energy usage line demonstrates a good trade off between the peeks when using a smart combination of simple single-hop and multi-hop network setups. The network lifetime is a lot higher compared to the single-hop or multi-hop approaches of figure B.5.

In the full binary tree structure of this example, we would have to relay for 8 nodes on level 4 if we do not want to add extra nodes. If no additional relay devices are used, data of up to 16 devices at level 1 can be relayed by the nodes of level 4 and up to 14 devices by the nodes of level 5.

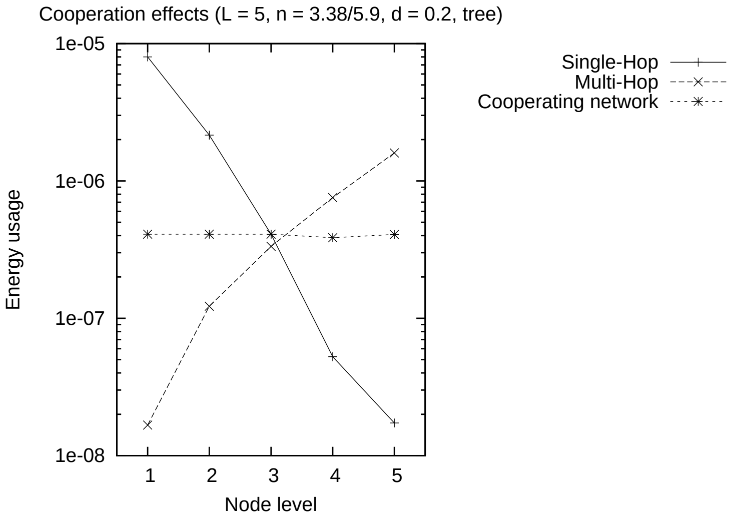


Figure B.7: Energy usage when cooperating with inter-node distance 20 cm

B.6 Future Work

The presented solutions are only a first step toward a highly energy efficient BASN. Based on the results carried out from this paper, we will develop new and adapt existing communication protocols for body area sensor networks. They will be based on the cooperation techniques, as proposed in this paper. Algorithms should be developed that can decide which node to forward to and that allow a node to communicate whether it is capable to cooperate. Existing cluster rotation techniques can be used as a starting point. Cooperation setup should be possible on-the-fly to facilitate the process of adding new sensors for end users. Another aspect to study is the optimal placement of relay nodes or the optimal number of nodes to cooperate with. An analytical approach to this problem can act as a benchmark for algorithmic solutions. Future research will also include an analysis of energy consumption in other frequency bands, next to the 2.4 GHz ISM band. Working in lower frequencies near the human body can result in a different performance.

B.7 Conclusions

In this paper we have studied the energy consumption in short-range networks experiencing high path loss, more specifically in body area sensor networks. Our results show that neither the classical single-hop approach nor multi-hop leads to a reasonable energy consumption. The high path loss has a large impact on the energy consumption when using single-hop on nodes far from the sink and hotspots appear near the sink when multi-hop is used. We have shown that using relay devices or a more cooperative approach can improve energy consumption largely, as this spreads out the transmission effort over the entire network.

B.8 Acknowledgments

This research is partly funded by the Fund for Scientific Research - Flanders (F.W.O.-V., Belgium) project G.0531.05, by The Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT - Vlaanderen) through the contract number 020152 and a PhD. grant for B. Latré and by the Flemish IBBT project IM3.

References

[1] C. Otto, A. Milenkovic, C. Sanders, and E. Jovanov. *System Architecture of a Wireless Body Area Sensor Network for Ubiquitous Health Monitoring*. Journal of Mobile Multimedia, 1(4):307–326, 2006.

[2] C. C. Y. Poon, Yuan-Ting Zhang, and Shu-Di Bao. *A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-health*. IEEE Communications Magazine, 44(4):73–81, April 2006.

[3] D. Cypher, N. Chevrollier, N. Montavont, and N. Golmie. *Prevailing over wires in healthcare environments: benefits and challenges*. IEEE Communications Magazine, 44(4):56–63, April 2006.

[4] P. J. Riu and K. R. Foster. *Heating of tissue by near-field exposure to a dipole: a modelanalysis*. IEEE Transactions on Biomedical Engineering, 46(8):911–917, August 1999.

[5] Q Tang, N. Tummala, S. K. S. Gupta, and L. Schwiebert. *Communication scheduling to minimize thermal effects of implanted biosensor networks in homogeneous tissue*. IEEE Transactions on Biomedical Engineering, 52(7):1285–1294, July 2005.

[6] International Commission on Non-ionizing Radiation Protection (ICNIRP). *Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz).* Health Physics, 74(4):494–522, apr 1998.

[7] *IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz.* 1999.

[8] S. K. S. Gupta, S. Lalwani, Y. Prakash, E. Elsharawy, and L. Schwiebert. *Towards a propagation model for wireless biomedical applications*. In Communications, 2003. ICC '03. IEEE International Conference on, volume 3, pages 1993–1997, May 2003.

[9] E. Reusens, W. Joseph, G. Vermeeren, and L. Martens. *On-body Measurements and Characterization of Wireless Communication Channel for Arm and Torso of Human*. In International Workshop on Wearable and Implantable Body Sensor Networks (BSN'07), pages 26–28, Aachen, March 2007.

[10] E. Reusens, W. Joseph, G. Vermeeren, L. Martens, B. Latré, B. Braem, C. Blondia, and I. Moerman. *Path-Loss Models for Wireless Communication Channel along Arm and Torso: Measurements and Simulations*. In IEEE AP-S Internation Symposium 2007, JUN 2007.

[11] A. Fort, J. Ryckaert, C. Desset, P. De Doncker, P. Wambacq, and L. Van Biesen. *Ultra-wideband channel model for communication around the human body*. IEEE Journal on Selected Areas in Communications, 24:927–933, April 2006.

[12] M. Welsh. *Exposing resource tradeoffs in region-based communication abstractions for sensor networks.* Computer Communication Review, 34(1):119–124, 2004.

[13] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. *Energy-efficient communication protocol for wireless microsensor networks*. In System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on, January 2000.

[14] J. C. Cano and P. Manzoni. *A performance comparison of energy consumption for Mobile Ad HocNetwork routing protocols*. In Modeling, Analysis and Simulation of Computer and Telecommunication Systems, 2000. Proceedings. 8th International Symposium on, pages 57–64, San Francisco, CA, USA, 2000.

[15] T. Falck, J. Espina, J. P. Ebert, and D. Dietterle. *BASUMA - the sixth sense for chronically ill patients*. In Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on, April 2006.

[16] B. Gyselinckx, R. Vullers, C. V. Hoof, J. Ryckaert, R. F. Yazicioglu, P. Fiorini, and V. Leonov. *Human++: Emerging Technology for Body Area Networks*. In Very Large Scale Integration, 2006 IFIP International Conference on, pages 175–180, October 2006.

[17] A. Milenkovic, C. Otto, and E. Jovanov. *Wireless sensor networks for personal health monitoring: Issues and an implementation*. Computer Communications, Wireless Sensor Networks and Wired/Wireless Internet Communications, 29(13-14):2521–2533, August 2006.

[18] T. Zasowski, F. Althaus, M. Stager, A. Wittneben, and G. Troster. *UWB for noninvasive wireless body area networks: channel measurements and results*. In Ultra Wideband Systems and Technologies, 2003 IEEE Conference on, pages 285–289, November 2003.

[19] B. Latré, G. Vermeeren, I. Moerman, L. Martens, and P. Demeester. *Networking and Propagation Issues in Body Area Networks*. In 11th Symposium on Communications and Vehicular Technology in the Benelux, SCVT 2004, November 2004.

[20] S. R. Saunders. *Antennas and propagation for wireless communication systems*. Wiley, West Sussex, 1999.

[21] Nordic, nRF 2401 datasheet [online] http://www.nordicsemi.com /index.cfm?obj=product&act=display&pro=64.

[22] Chipcon, CC2420 datasheet [online] http://focus.ti.com/docs/prod /folders/print/cc2420.html.

C

A Secure Cross-layer Protocol for Multi-hop Wireless Body Area Networks

D. Singelée[1], B. Latré[2], B. Braem[3], M. Peeters[4], M. De Soete[4], P. De Cleyn[3], B. Preneel[1], I. Moerman[2] and C. Blondia[3]

Accepted for presentation at the 7th International Conference on AD-HOC Networks & Wireless (ADHOCNOW 2008), France, September 2008

[1] Katholieke Universiteit Leuven – IBBT, ESAT – SCD – COSIC, Leuven, Belgium
[2] Ghent University – IBBT, Dept. of Information Technology, IBCN research group, Gent, Belgium
[3] University of Antwerp – IBBT, Dept. of Mathematics and Computer Science, PATS research group, Antwerp, Belgium
[4] NXP Semiconductors, Competence Center System Security & DRM, Leuven, Belgium

Abstract *The development of Wireless Body Area Networks (WBANs) for wireless sensing and monitoring of a person's vital functions, is an enabler in providing better personal health care whilst enhancing the quality of life. A critical factor in the acceptance of WBANs is providing appropriate security and privacy protection of the wireless communication. This paper first describes a general health care platform and pinpoints the security challenges and requirements. Further it proposes and analyzes the CICADA-S protocol, a secure cross-layer protocol for*

WBANs. It is an extension of CICADA, which is a cross-layer protocol that handles both medium access and the routing of data in WBANs. The CICADA-S protocol is the first integrated solution that copes with threats that occur in this mobile medical monitoring scenario. It is shown that the integration of key management and secure, privacy preserving communication techniques within the CICADA-S protocol has low impact on the power consumption and throughput.

C.1 Introduction

Recent progress in wireless sensing and monitoring, and the development of small wearable or implantable biosensors, have led to the use of Wireless Body Area Networks (WBANs). The research on communication within a WBAN is still in its early stages. Only few protocols designed specifically for multi-hop communication in WBANs exist. They try to minimize the thermal effects of the implanted devices by balancing the traffic over the network [1] or by forming clusters [2, 3] or a tree network [4].

Wireless Body Area Networks can be seen as an enabling technology for mobile health care [5]. Medical readings from sensors on the body are sent to servers at the hospital or medical centers where the data can be analyzed by professionals. These systems reduce the enormous costs associated to ambulant patients in hospitals as monitoring can take place even at home in real-time and over a longer period.

In this paper, we propose and analyze CICADA-S, a secure protocol for WBANs. It is based on an existing multi-hop protocol for WBANs, called CICADA [4]. This is a cross-layer protocol that sets up a data gathering tree in a reliable manner, offering low delay and high energy efficiency. The communication of health related information between sensors in a WBAN and over the Internet to servers is strictly private and confidential and should therefore be encrypted to protect the patient's privacy. Furthermore, the medical staff who collects the data must be confident that the data is not tampered with, and indeed originates from that patient.

The CICADA-S protocol is designed within the scope of the IBBT IM3-project (Interactive Mobile Medical Monitoring), which focuses on the research and implementation of a wearable system for health monitoring [6]. Patient data is collected using a WBAN and analyzed at the gateway (also called medical hub) worn by the patient. If an event (e.g., heart rhythm problems) is detected, a signal is sent to a health care practitioner who can view and analyze the patient data remotely.

The remainder of this paper is organized as follows. Section C.2 gives an overview of related work. The general architecture and the necessary security assumptions are described in section C.3. A short description of CICADA is given, followed by the integration of the security mechanisms in the protocol and a description of the key management aspects in section C.4. The analysis of the integration in terms of performance overhead and the security properties are dealt with in section C.5. Finally, section C.6 provides a final conclusion on the paper.

C.2 Related Work

Security is essential for broad acceptance and further growth of Wireless Sensor Networks. These networks pose unique challenges as security techniques used in traditional networks cannot be directly applied. Indeed, to make sensor networks economically viable, sensor devices should be limited in their energy consumption, computation, and communication capabilities. Since most of the existing security mechanisms have major drawbacks in that respect, new ideas are needed to address these requirements in an appropriate way [7].

One of the most crucial components to support the security architecture of a Wireless Sensor Network is its key management. During the last years, a number of pairwise key establishment schemes have been proposed. Zhou and Haas propose to secure ad-hoc networks using asymmetric cryptography [8]. They use threshold cryptography to distribute trust among a set of servers. This scheme achieves a high level of security, but is too energy consuming to be used in practice in a Wireless Sensor Network. Eschenauer and Gligor introduce a key management scheme for distributed sensor networks [9]. It relies on probabilistic key sharing among the nodes of a random graph. Perrig et al. present SPINS, a suite of security building blocks optimized for resource-constrained environments and wireless communication [10]. It has two secure building blocks: SNEP and μTESLA. SNEP provides data confidentiality, two-party data authentication and data freshness, while μTESLA offers authenticated broadcast in constrained environments.

The security mechanisms employed in Wireless Sensor Networks do generally not offer the best solutions to be used in Wireless Body Area Networks for the latter have specific features that should be taken into account when designing the security architecture. The number of sensors on the human body, and the range between the different nodes, is typically quite limited. Furthermore, the sensors deployed in a WBAN are under surveillance of the person carrying these devices. This means that it is difficult for an attacker to physically access the nodes without this being detected. When designing security protocols for WBANs, these characteristics should be taken into account in order to define optimized solutions with respect to the available resources in this specific environment.

Although providing adequate security is a crucial factor in the acceptance of WBANs, little research has been done in this specific field [11]. In [12] an algorithm based on biometric data is described that can be employed to ensure the authenticity, confidentiality and integrity of the data transmission between the personal device and all the other nodes. Biometrics is a technique commonly known as the automatic identification and verification of an individual by his or her physiological characteristics. Another method is presented in [13] where body-coupled communication (BCC) is used to associate new sensors in a WBAN.

None of the current protocols offer a solution where appropriate security mechanisms are incorporated into the communication protocol while addressing the lifecycle of the sensors. Further, security and privacy protection mechanisms use a significant part of the available resources and should therefore be energy efficient and lightweight. The mechanisms proposed in this paper aim to cover these challenges.

C.3 Architecture

C.3.1 General Overview

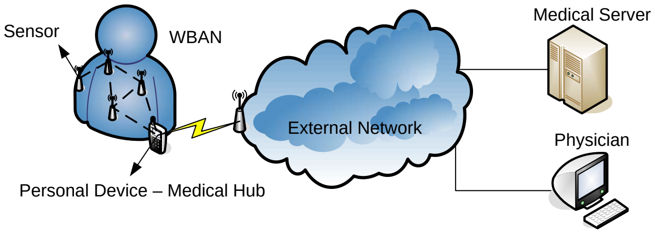


Figure C.1: General overview of the IM3 health care architecture.

Fig. C.1 shows the health care architecture used by the IM3 project. There are three main components: the Wireless Body Area Network (WBAN), the external network and the back-end server. In this scenario, the WBAN contains several sensors that measure medical data such as ECG, body movement etc. These sensors send their measurements, directly or via several hops, to the gateway. Each WBAN (and hence every patient) has its unique gateway. In other words, the sensors shall only send their data to the unique gateway they are linked with and this needs to be enforced by specific security mechanisms. The gateway processes the medical data, and sends the result via the external network to the back-end server at the hospital, where it can be observed and analyzed by medical staff.

Although the architecture was originally designed for and is fully adapted to a medical environment, it may also be used in other applications. Indeed, as long as the (security) relations between the different devices remain valid, the protocol remains applicable, which increases the generality of our solution. In the remainder of this paper, the medical scenario will be further used to explain the architecture and the secure cross-layer protocol for multi-hop WBANs.

C.3.2 Security Assumptions

This section aims to address the security of the entire system, and the WBAN in particular.

The most security critical device in the entire architecture is the back-end server. This server, which is managed by the hospital or medical center, will receive the medical data sent by all active WBANs. It is assumed that this server is physically protected (e.g., put in a secure place in the hospital where it can not be stolen or tampered with), and that an adequate access control system is implemented (i.e. only authorized medical personnel has (partial) access to the server through appropriate identification/authentication mechanisms). The back-end server is considered to be a trusted third party, which means that it is known and trusted by all other devices in the network after a successful authentication.

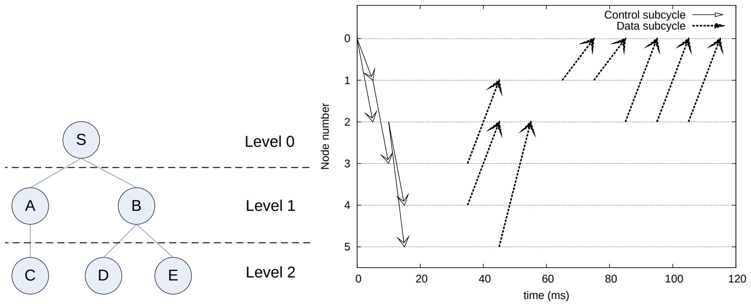
Since potentially security critical data will be transferred through the external network, end-to-end security between the gateway and the back-end server is required. For efficiency reasons, it is assumed that both devices share a symmetric session key to secure their communication. This symmetric session key can be manually installed (e.g., pre-installed during manufacturing), or (preferably) established via a symmetric key establishment protocol. The description of such protocols can be found in the ISO 9798–2 standard, and is out of scope of this article. The symmetric session key is updated regularly. The end-to-end channel between gateway and back-end server should also be anonymized using temporary pseudonyms. This avoids privacy problems like (location) tracking. In the remainder of the paper, it is assumed that the secure end-to-end channel between gateway and back-end server is already established after a successful mutual authentication. As mentioned before, each gateway belongs to a specific WBAN (i.e. a patient, who is carrying this device). To enforce this, the gateway is registered in advance at the back-end server.

It is assumed that it is impossible to alter or read the memory of a (securely initialized) node that is put on the patient's body, or to modify the behavior of a node without this being detected. This is not a strong assumption, since the patient is carrying the nodes on its body, and an attacker is not able to access the nodes without this being detected. It is also assumed that the attacker has no access to the sensors that yet have to be securely initialized (e.g., because they are stored in a safe place). However, an attacker can put a malicious node in the presence of a WBAN, and try to join the network. He can also eavesdrop on all data transmitted in the WBAN, and insert/delete/modify (malicious) data into the network. The attacker is hence assumed to be active.

C.4 Protocol Design

C.4.1 CICADA

CICADA is a cross-layer protocol as it handles both medium access and the routing of data [4]. The protocol sets up a spanning tree in a distributed manner, which is subsequently used to guarantee collision free access to the medium and to route data toward the gateway. The time axis is divided in slots grouped in cycles, to lower the interference and avoid idle listening. Slot assignment is done in a distributed way where each node informs its children when they are allowed to send their data using a SCHEME. Slot synchronization is possible because a node knows the length of each cycle. During a cycle, a node is allowed to send all of its data to its parent node. CICADA is designed in such a way that all packets arrive at the source in only one cycle. Routing itself is not complicated in CICADA anyway as data packets are routed up the tree which is set up to control the medium access, no special control packets are needed.



(a) Sample topology

(b) Packet streams in this network. Notice the down- and upstream traffic

Figure C.2: Communication in CICADA for a sample network of 5 nodes

A cycle is divided in a control subcycle consisting of control slots, and a data subcycle consisting of data slots. The former is used to broadcast a SCHEME message from parent to child, i.e. to let the children know when they are allowed to send in the data subcycle. In the data subcycle, data is forwarded from the nodes to the gateway. In each data subcycle, a contention slot is included to allow nodes to join the tree. New children hear the SCHEME message of the desired parent and send a JOIN-REQUEST message in the contention slot. When the parent hears the JOIN-REQUEST message, it will include the node in the next cycle. Each node will send at least two packets per cycle: a data packet or HELLO packet (if no data is sent) and a SCHEME packet. If a parent does not receive a packet from a child for N or more consecutive cycles, the parent will consider the child to be lost. If a child does not receive packets from its parent for N or more consecutive cycles, the child will assume that the parent is gone and will try to join another node. An example of communication in CICADA is given in Fig. C.2, for a network of 5 nodes. The control and data subcycles can be seen clearly.

A node informs its parent node of the number of slots it needs to send its own data and forward data coming from its children, by calculating two parameters: α and β. The former gives the number of slots needed for sending data (including forwarded data) to its parent, the latter gives the number of slots the node has to wait until it has received all data from its children. Based on the α and β from its children, a node can calculate the slot allocation for the next cycle.

C.4.2 CICADA-S

The CICADA protocol, as described in the previous section, does not guarantee any form of security and privacy. Unauthorized nodes can easily join the WBAN, and all communication in the network is sent in plain text and is not integrity protected. The fixed identity of the sensors is not kept confidential, and can hence be used to track sensors (and patients carrying these sensors). To counter these problems, appropriate security mechanisms have to be added to the CICADA protocol. The result is the CICADA-S protocol, the secure version of the CICADA protocol.

From a security point of view, there are four main states which take place during the lifetime of a sensor: the secure initialization phase, the sensor (re)joining the WBAN, a key update procedure in the WBAN, and the sensor leaving the WBAN. The security mechanisms used in these phases and their integration into the CICADA-S protocol, based on the results of [6], will now be described.

C.4.2.1 Secure Initialization Phase:

Initially, each sensor has to be securely initialized by the back-end server before it can join the WBAN in a later stage. During this initialization phase, the sensor and the back-end server will agree on a shared symmetric key. This can be done via asymmetric cryptographic techniques, but this is typically too energy (and computation) consuming for a regular sensor. Another way of establishing a shared key, is by using a private and authentic out-of-band channel. Such a channel is typically cheap to setup. It has the interesting property that all data transmitted on the channel remains confidential for eavesdroppers, and that the integrity and authenticity is protected too. A private and authentic channel can be created in several ways, depending on the exact hardware and (physical) characteristics of the sensors. It can be established by connecting the sensor directly to the back-end server, via an extra electrical contact available on both devices. Other techniques to create such a secure out-of-band channel is by employing distance bounding protocols, by having the user manually enter the data on both devices etc. More information on these and other techniques to establish a private and authentic out-of-band channel can be found in the literature [14–16].

Let us assume that sensor *A* has to be initialized. The data transfer via the secure out-of-band channel takes place in two steps. First, the sensor sends its fixed identity to the back-end server. This can be done explicitly or implicitly (the identity of the sensor can be implicitly known because of the specific characteristics of the out-of-band channel). In the second step of the protocol, the back-end server generates a random secret key (k_A), and sends this key securely to the sensor. The sensor and the back-end server store this secret key in their memory. The key is (conceptually) composed out of 2 subkeys: the encryption key k_{A_encr} and the integrity key k_{A_int}. Note that each new node is assigned a new and unique secret key.

Each sensor *i* is also assigned a unique counter CTR_i, which is initialized to 0 and stored in the sensor's memory. The value of this counter is included in all key management messages, and is used to avoid replay attacks and assure freshness. Every time the counter is used, the value gets incremented by 1.

C.4.2.2 Sensor (re)joining the WBAN:

After the initialization procedure, the sensor is ready to be put on the patient's body. It will detect the WBAN, and start the join procedure, which will now be discussed.

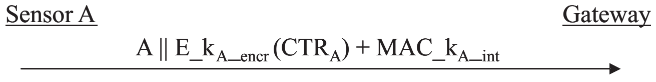


Figure C.3: Secure JOIN-REQUEST originating from sensor *A*.

When the sensor (with fixed identity A) hears the SCHEME of the desired parent, it sends a secure JOIN-REQUEST message, as shown in Fig. C.3, in the contention slot. This message is forwarded to the gateway. It is basically a HELLO message containing the unique (global) identity of the sensor and the value of its unique counter CTR_A. The counter is encrypted for privacy reasons (since it is used in all key management messages). The gateway stores (and updates) this value of the counter. The integrity and authenticity of the entire secure JOIN-REQUEST message is protected by a message authentication code (*MAC*) [17], computed with the key k_{A_int}.

When the gateway receives the secure JOIN-REQUEST message of sensor *A*, it forwards this request to the back-end server via the secure end-to-end channel. This triggers a protocol in which the key k_A is securely transported from the back-end server to the gateway. More information on how to accomplish this, can be found in the ISO 9798–2 standard [18]. In some scenarios, and this is often the case in a medical environment, it is known in advance (e.g., already during the initialization procedure) in which WBAN the sensor will be deployed. In this case, the back-end server can already transport the key k_A to the correct gateway, and does not have to wait until it receives the secure JOIN-REQUEST message. This makes the join procedure faster. In the case a sensor leaves the network, and (not much) later rejoins it, the gateway may still have the key k_A in its memory and does not have to forward the request to the back-end server. From the moment the gateway has access to the key, it can check the validity of the JOIN-REQUEST by verifying the message authentication code, and in case of a rejoin, also the value of the counter CTR_A (the new value should be higher than the current value shared by sensor and gateway). If this verification is successful, the sensor is allowed to join the WBAN and is assigned a temporary identity $localID_A$. This temporary identity, which is chosen by the gateway, is established in order to preserve the privacy. It is only unique within the environment of the WBAN. Other networks can reuse the same identifier. Since the bitlength of such a local identifier can be smaller than the full identity of the sensor (*A*), it also improves the efficiency. A joining sensor in the WBAN is informed about its temporary identity during the key transport procedure, which takes place immediately after the approval of the secure JOIN-REQUEST message.

C.4.2.3 Key Update Procedure in the WBAN:

Except for the key management messages, the data traveling in the WBAN consists of schemes sent during the control subcycle, and medical data sent during the data subcycle from the sensors to the gateway. The former is only integrity protected (to allow a new node to inform itself about the contention slot), while the latter is both integrity protected and encrypted. All these operations are performed by employing a secret group key s, that is shared between all the sensors in the WBAN. Every time a node joins or leaves the network, the group key is updated in order to avoid an attacker recovering the key. Even when the topology of the network remains constant for a long time, the group key should still be updated at regular intervals. The exact period is determined by the cryptographic strength of the encryption and integrity algorithms used to protect the data in the WBAN, and the length of the key. We will briefly come back to this in section. C.5.1.

The update process works as follows. First, the gateway randomly generates a new group key s. Next, it performs a secure key transport procedure with all the nodes in the WBAN, as shown in Fig. C.4. The gateway constructs a key update message, unique for every sensor, which contains the encrypted value of the updated group key s. For each node i, the message also contains the new value of the counter CTR_i (which is the current value of the counter incremented by 1), in order to avoid replay attacks, and the local identifier $localID_i$. The authenticity and the integrity of the message is protected by a message authentication code. Nodes that have been excluded from the WBAN, can not decrypt the key transport messages anymore, and are hence not able to obtain the new group key s.

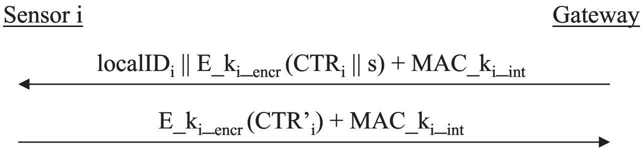

Sensor i
Gateway
localID$_i$ || E_k$_{i\_encr}$(CTR$_i$ || s) + MAC_k$_{i\_int}$
E_k$_{i\_encr}$(CTR'$_i$) + MAC_k$_{i\_int}$

Figure C.4: Secure key transport to all the sensors in the WBAN.

The key update message is uniquely constructed for every sensor, and forwarded from the gateway to the correct node during the control subcycle. Each node takes the message containing its local identifier, checks the validity of the message (by verifying the value of the counter and the message authentication code) and decrypts the encrypted part in order to recover the new value of the group key s. It also forwards all other key update messages to its children, who perform the same procedure. A new joining node *A* does not yet know its local identifier $localID_A$, and therefore has to check the message authentication code (and the counter) of all the key update messages using its key k_{A_int} until the test succeeds. This only has to be done once, and is easily feasible since computing a message authentication code can be done very efficiently. The joining sensor stores its local identifier $localID_A$ in its memory, and recovers the group key s from the encrypted part of the key update message. Finally, all sensors send a secure acknowledgement back to the gateway during the next data subcycle, to inform that they received the key well. This key confirmation message only contains the encrypted value of the updated counter CTR_i, concatenated with a message authentication code. After having received the key confirmation message, the gateway knows it can definitively update the group key. When a node does not send its key confirmation message within a certain period, e.g., because it did not receive the new group key s due to packet loss, the gateway retransmits the key transport message to that particular node.

C.4.2.4 Sensor Leaving the WBAN:

When a node detects that a particular sensor *A* is not part anymore of the WBAN, it forwards this information to the gateway. This automatically triggers a group key update procedure. This has to be done in order to avoid that an attacker stealing a sensor from the network, would be able to read or modify the data in the WBAN. After a certain interval (or even immediately, depending on the policy), the gateway deletes the key k_A and the identifier $localID_A$ from its memory. If the medical staff removes sensor *A* from the patient, or if the sensor is reported lost or stolen, the key k_A should also be deleted from the memory of the back-end server. This way, the sensor can not rejoin any network anymore in a later stage, until it has been securely reinitialized by the back-end server.

C.5 Analysis

C.5.1 Performance Evaluation

The addition of these security mechanisms to CICADA undoubtedly influences the performance as it leads to an increased overhead and higher delay. The exact impact strongly depends on the choice of the cryptographic algorithms that are deployed in the WBAN, and it is hence difficult to formulate results that are generally applicable. That is why a worst case analysis will be given, in which we assume that a secure block cipher, such as the Advanced Encryption Standard (AES) [19], is employed in an authenticated encryption mode (e.g., CCM or GCM mode of operation). The numbers used below are based on the guidelines of the National Institute of Standards and Technology (NIST) [20, 21]. In practice, it would be better to employ a low-cost encryption and integrity algorithm, which has a slightly lower security level, but is more efficient.

The combined encryption and authentication algorithm uses a symmetric key of 16 bytes (the group key *s* or the shared key k_i). The output of this method are encrypted blocks of 16 bytes, and a message authentication code of at least 8 bytes. Furthermore, the unique hardware address of the sensor is assumed to be 6 bytes (e.g., as in Bluetooth), and a counter of 4 bytes is employed to avoid replay attacks. Note that encrypting the counter results in an encryption block of 16 bytes. Using these parameters offers a high level of security as long as the keys are updated regularly, which depends on the strength of the cryptographic algorithm that is being used. E.g., when AES is used in the GCM mode of operation, the group key *s* should be updated at least at every 2^{32}th invocation of the encryption algorithm [21]. In this section, we will now briefly discuss the (worst case) impact of the security mechanisms on the CICADA protocol, using the numbers stated above.

In the (re)joining phase, additional information is sent to the gateway in the JOIN-REQUEST message. The original CICADA-message only contains $localID_A$ and $localID_P$ (i.e. the local ID of node A joining the network and the local ID of the desired parent P respectively). The length of these IDs is 1 byte, which is sufficient for a WBAN. In CICADA-S the unique hardware address of the sensor is sent, together with the encrypted synchronized counter and a message authentication code. The length of the JOIN-REQUEST message thus is longer, but still only 30 bytes. As this information is sent in a contention slot with fixed size, this will not influence the throughput of the system. However, this secure JOIN-REQUEST message needs to be forwarded to the gateway. As the contention slot of a node is in the beginning of a data subcycle, the message can be sent to the gateway directly. E.g., the JOIN-REQUEST message can be piggybacked on a data packet that is sent to the gateway. As the length of the message is small, this may not influence the overall throughput significantly. The number of bytes that can be sent in one slot depends on the size of the slot and the raw bit rate of the radio technology used. If the number of bytes in the data packet and the secure JOIN-REQUEST message is too large, the slot size will have to be altered. This will lower the throughput of the network. A better solution is to send the JOIN-REQUEST message in a separate data slot. This will hardly impact the throughput of the network. If the key is already present at the gateway, the gateway can immediately start the key update procedure. If not, the gateway has to wait for a response from the back-end server. This will add extra delay to the joining procedure.

In the key update procedure, the gateway sends a new key to all the nodes in the control subcycle. This message contains $localID_A$, the new key group key s concatenated with an increased counter (both encrypted), and a message authentication code. For each node, this is an additional 41 bytes. Due to the broadcast mechanism in the control subcycle, these messages all need to be broadcasted by every node sending its SCHEME in the control subcycle. This will lead to a larger slot length in the control subcycle, and subsequently a lower throughput. In CICADA, the slot length in the control subcycle is smaller than the data slot length as the SCHEME-messages sent in the control subcycle are very short. The slot length can be up to ten times smaller. This improves the energy throughput of CICADA. As the key is only updated after several cycles, we opt to change the control slot dynamically. When the key is updated, the control slot length has the same length as the data slot. At any other time, the control slot has its shorter length. When the key is about to be updated, the gateway broadcasts a warning in the previous cycle by setting a bit in the header. The nodes receive this warning and adapt their control slot lengths for one cycle.

When a node leaves the network or is no longer attached to it, the (former) parent node sends a message to the gateway. This can be added to a data packet and will not influence the throughput.

It is very important to note that the key management messages are sent rarely (only when a node (re)joins the network, or when the group key has to be updated), and hardly affect the global throughput in the network. Most data traveling in the WBAN is medical data, sent by the sensors to the gateway. These messages are protected by employing the group key s. The data is encrypted in blocks of 16 bytes, and a message authentication code of 8 bytes is added. The SCHEME packets sent during the control subcycle are not encrypted, but integrity protected.

For both types of data, the length of the messages is hardly influenced. Overall, the security mechanisms will have a minor impact on the performance of CICADA-S.

C.5.2 Security Claims

One of the design goals of the CICADA-S protocol is to secure the wireless communication in the WBAN while preserving privacy. The most interesting security properties of our protocol will now be briefly discussed (without formal proof). It has to be stressed that the following statements are based on the assumptions stated in section C.3.2, and that all devices in the network, including the attacker, are computationally bounded.

- The CICADA-S protocol provides forward security. A node that leaves the network can not successfully read/modify/insert/delete data in the WBAN, since the group key s is always updated in case the topology of the network changes.

- Nodes that are not securely initialized, can not join the WBAN. Only nodes that share a symmetric key with the back-end server, can construct a valid secure JOIN-REQUEST message, which is needed to join the WBAN.

- Since the group key is transported in an encrypted format from the gateway to the nodes in the WBAN, it is practically not feasible for an eavesdropper to recover the key. Only an attacker that can break the encryption scheme used to protect data in the WBAN, is able to find the group key s.

- The CICADA-S protocol offers key confirmation, which is important for security and performance reasons. After receiving the new group key s, a node sends a key confirmation message to the gateway, to inform that the key was received well. This avoids certain Denial-of-Service attacks (e.g., blocking key update messages). Due to packet loss and bit errors, key confirmation is also an important and necessary property of network protocols for wireless media.

- A sensor that is a member of a WBAN can not join another WBAN at the same time. The second secure JOIN-REQUEST message sent by the sensor will be refused by the back-end server, because this device will detect that the sensor already belongs to another network.

- Nodes that are part of a particular WBAN, are not able to read, modify, insert or delete encrypted data in other WBANs without this being detected, since these other networks do not share the same group key s.

- Since the confidentiality and integrity of data traveling in the WBAN is cryptographically protected, a device that does not possess the group key will not succeed in decrypting the enciphered communication in the WBAN, nor successfully modifying/inserting/deleting data into the network without this being detected.

- Replay attacks are detected because of the use of the synchronized counter, that is shared between sensor and gateway.
- Location privacy has been taken into account during the design of the CICADA-S protocol. The communication between gateway and back-end server is assumed to be completely secured (end-to-end) and anonymized. Using the data in the WBAN to trace a patient is also not possible, because it only contains local identifiers, and these are not unique across WBANs. Only in the first message of the join procedure, the exact identity of the sensor is exposed. It is however not used in the other key management messages. Neither is it possible to link other messages to the initial key management message of the join procedure (since the synchronized counter is encrypted). As a result, the data in the WBAN can not be used to trace patients.

C.6 Conclusion

Wireless Body Area Networks are an enabling technology for mobile health care. These systems reduce the enormous costs associated to patients in hospitals as monitoring can take place even at home in real-time and over a longer period. A critical factor in the acceptance of WBANs is the provision of appropriate security and privacy protection of the wireless communication medium. The data traveling between the sensors should be kept confidential and integrity protected. Certainly in the mobile monitoring scenario, this is of uttermost importance.

In this paper we have presented CICADA-S, a security enabled cross-layer multi-hop protocol for Wireless Body Area Networks. It is a secure extension of the CICADA protocol, and was designed within the scope of the IM3-project (Interactive Mobile Medical Monitoring), which focuses on the research and implementation of a wearable system for health monitoring. The CICADA-S protocol is the first integrated solution to cope with the threats of interactive mobile monitoring and the life cycle of the sensors. It combines key management and secure privacy preserving communication techniques. We have presented the main security properties of CICADA-S, and shown that the addition of security mechanisms to the CICADA-S protocol has low impact on the power consumption and throughput. The security mechanisms integrated in the protocol are simple, yet very effective. The CICADA-S protocol can be implemented on today's devices as it only requires low-cost and minimal hardware changes.

The authors strongly believe that adding sufficient security mechanisms to Wireless Body Area Networks will work as a trigger in the acceptance of this technology for health care purposes.

Acknowledgments.

This work is partially funded by a research grant of the Katholieke Universiteit Leuven for D. Singelée, by the Concerted Research Action (GOA) Ambiorics 2005/11 of the Flemish Government, by the IAP Programme P6/26 BCRYPT of the Belgian State (Belgian Science Policy), by the Fund for Scientific Research — Flanders (F.W.O.-V.,Belgium) project G.0531.05 (FWO-BAN) and by the Flemish IBBT project IM3.

References

[1] D. Takahashi, Y. Xiao, and F. Hu. *LTRT: Least Total-Route Temperature Routing for Embedded Biomedical Sensor Networks*. In IEEE Globecom 2007, November 2007.

[2] M. Moh, B. J. Culpepper, Lan Dung, Teng-Sheng Moh, T. Hamada, and Ching-Fong Su. *On data gathering protocols for in-body biomedical sensor networks*. In Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE, volume 5, November/December 2005.

[3] A. G. Ruzzelli, R. Jurdak, G. M.P O'Hare, and P. Van Der Stok. *Energy-efficient multi-hop medical sensor networking*. In HealthNet '07: Proceedings of the 1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments, pages 37–42, New York, NY, USA, 2007. ACM.

[4] B. Latré, B.Braem, I.Moerman, C. Blondia, E. Reusens, W. Joseph, and P. Demeester. *A Low-Delay Protocol for Multihop Wireless Body Area Networks*. In Mobile and Ubiquitous Systems: Networking & Services, 2007 4th Annual International Conference on, Philadelphia, PA, USA, August 2007.

[5] C. Otto, A. Milenkovic, C. Sanders, and E. Jovanov. *System Architecture of a Wireless Body Area Sensor Network for Ubiquitous Health Monitoring*. Journal of Mobile Multimedia, 1(4):307–326, 2006.

[6] IBBT IM3-project [online] http://projects.ibbt.be/im3.

[7] A. Perrig, J. Stankovic, and D. Wagner. *Security in Wireless Sensor Networks*. Communications of the ACM, 47(6):53–57, June 2004.

[8] L. Zhou and Z. J. Haas. *Securing Ad Hoc Networks*. IEEE Network, 13(6):24–30, November/December 1999.

[9] Laurent Eschenauer and Virgil D. Gligor. *A key-management scheme for distributed sensor networks*. In CCS '02: Proceedings of the 9th ACM conference on Computer and communications security, pages 41–47, New York, NY, USA, 2002. ACM.

[10] A. Perrig, R. Szewczyk, V. Wen, D.E. Culler, and J.D. Tygar. *SPINS: Security Protocols for Sensor Networks*. In Mobile Computing and Networking, pages 189–199, 2001.

[11] H. Baldus, K. Klabunde, and G. Msch. *Reliable Set-Up of Medical Body-Sensor Networks.* In Holger Karl, Andreas Willig, and Adam Wolisz, editors, EWSN, volume 2920 of *Lecture Notes in Computer Science*, pages 353–363. Springer, 2004.

[12] C. C. Y. Poon, Yuan-Ting Zhang, and Shu-Di Bao. *A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-health*. IEEE Communications Magazine, 44(4):73–81, April 2006.

[13] T. Falck, H. Baldus, J. Espina, and K. Klabunde. *Plug 'n play simplicity for wireless medical body sensors*. Mob. Netw. Appl., 12(2-3):143–153, 2007.

[14] C. Gehrmann, C. Mitchell, and K. Nyberg. *Manual Authentication for Wireless Devices*. RSA Cryptobytes, 7(1):29–37, 2004.

[15] D. Singelée and B. Preneel. *Key Establishment Using Secure Distance Bounding Protocols*. In First Workshop on the Security and Privacy of Emerging Ubiquitous Communication Systems, SPEUCS 2007, Philadelphia, PA, USA, August 2007. IEEE.

[16] F. Stajano and R. Anderson. *The Resurrecting Duckling: Security Issues in Ad–Hoc Wireless Networks*. In Proceedings of the 7th International Workshop on Security Protocols, Lecture Notes in Computer Science, LNCS 1796, pages 172–182. Springer-Verlag, 1999.

[17] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1996.

[18] ISO/IEC 9798-2. *Information technology – Security techniques – Entity authentication – Part 2: Mechanisms using symmetric encipherment algorithms*. ISO, 1999.

[19] J. Daemen and V. Rijmen. *The Design of Rijndael – AES - The Advanced Encryption Standard*. Springer-Verlag, 2002.

[20] NIST Special Publication 800-38C. *Recommendation for Block Cipher Modes of Operation The CCM Mode for Authentication and Confidentiality*. U.S. DoC/NIST. Available at `http://csrc.nist.gov/publications/`, May 2004.

[21] NIST Special Publication 800-38D. *Recommendation for Block Cipher Modes of Operation Galois/Counter Mode (GCM) and GMAC*. U.S. DoC/NIST. Available at `http://csrc.nist.gov/publications/`, November 2007.