
UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola Tècnica Superior d'Enginyeria
de Telecomunicació de Barcelona


# Coarse-grained Localization of In-body Energy-harvesting Nanonodes

Degree Thesis submitted to the Faculty of Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona

Universitat Politècnica de Catalunya
by

Arnau Brosa Lopez

In partial fulfillment
of the requirements for the degree in
**Telecommunications Technologies and Services Engineering**

Advisors : Sergi Abadal and Filip Lemic
Barcelona, May 2023

# Contents

# List of Figures



# List of Tables

# Abstract

Nanoscale devices with Terahertz (THz) wireless communication capabilities are envisioned for sensing and actuation-based applications within human bloodstreams. These devices detect biomarkers, enable targeted drug delivery, and improve precision diagnostics. The introduction of flow-guided nanoscale localization utilizes THz-based communication to establish communication between nanonodes and anchors. This approach is envisaged to accurately locate regions where events occur by using the nanodevice's circulation duration in the bloodstream. This enables precise identification of disease biomarkers, viruses, and bacteria, facilitating targeted intervention and early detection of health conditions.

To avoid the pitfalls encountered in benchmarking and standardizing traditional indoor localization, this work presents a workflow for standardized performance evaluation of flow-guided nanoscale localization. The workflow is implemented in the form of an open source simulator, considering nanodevice mobility, in-body THz communication with on-body anchors, and energy-related constraints. The simulator is able to generate raw data that can be used to streamline different flow-guided localization solutions and establish standardized performance benchmarks.

The evaluation is performed in the form of a design space exploration. The results indicate that the proposed workflow and the simulator can be utilized for capturing the performance of flow-guided localization approaches in a way that allows objective comparison with other approaches serving as the foundation for standardized evaluation of future solutions.

# Resumen

Dispositivos a escala nanométrica con capacidades de comunicación inalámbrica en Terahertz (THz) aspiran a implementarse para aplicaciones de detección dentro del torrente sanguíneo humano. Estos dispositivos detectan biomarcadores, permiten la administración dirigida de medicamentos y mejoran el diagnóstico precoz. La introducción de la localización nanométrica guiada por flujo utiliza la comunicación basada en THz para establecer la comunicación entre nanonodos y anclas. Se espera que este enfoque localice con precisión las regiones donde ocurren los eventos mediante el tiempo de circulación del nanodispositivo en el torrente sanguíneo. Esto facilita la identificación precisa de biomarcadores de enfermedades, virus y bacterias, lo que permite una intervención dirigida y la detección temprana de diversas condiciones de salud.

Para evitar los desafíos encontrados en la evaluación y estandarización de la localización tradicional, este trabajo presenta un flujo de trabajo para la evaluación del rendimiento estandarizado de la localización nanométrica guiada por flujo. El flujo de trabajo se implementa en forma de un simulador de código abierto, teniendo en cuenta la movilidad del nanodispositivo, la comunicación THz en el cuerpo con anclas externas y las restricciones relacionadas con la energía. El simulador puede generar datos que se pueden utilizar para optimizar diferentes soluciones de localización y establecer puntos de referencia de rendimiento estandarizados.

La evaluación se realiza mediante una exploración del espacio de diseño. Los resultados indican que el flujo de trabajo propuesto y el simulador se pueden utilizar para capturar el rendimiento de los enfoques de localización guiados por flujo de manera que permita una comparación objetiva con otros enfoques, sentando así las bases para la evaluación estandarizada de soluciones futuras.

# Resum

Els dispositius a nanoescala amb capacitats de comunicació sense fils en Terahertz (THz) aspiren a implementar-se en aplicacions basades en detecció i actuació dins del torrent sanguini humà. Aquests dispositius detecten biomarcadors, permeten el lliurament precís de fàrmacs i proporcionen un diagnòstic precoç. La introducció de la localització a nanoescala guiada per flux utilitza la comunicació basada en THz per establir la comunicació entre nanonodes i ancoratges. Aquest enfocament aspira a localitzar amb precisió les regions on es produeixen els esdeveniments utilitzant el temps de circulació del nanodispositiu a cada una de les regions. Això permet la identificació precisa de biomarcadors de malalties, virus i bacteris, donant lloc a una intervenció dirigida i a la detecció precoç de diverses condicions de salut.

Per evitar els inconvenients que sorgeixen en les primeres etapes d'investigació, aquest treball presenta un flux de treball per a l'avaluació estandarditzada del rendiment de la localització a nanoescala guiada per flux. El flux de treball s'implementa en forma d'un simulador de codi obert, tenint en compte la mobilitat del nanodispositiu, la comunicació THz dins del cos amb els ancoratges situats a la superfície del cos i les limitacions relacionades amb l'energia. El simulador és capaç de generar dades que després de ser processades, ens permeten obtenir una avaluació estandaritzada del sistema .

L'avaluació es va realitzar en forma d'exploració espacial de disseny. Els resultats indiquen que el flux de treball proposat i el simulador es poden utilitzar per capturar el rendiment de la solució implementada d'una manera que permet la comparació objectiva amb altres enfocaments, servint d'aquesta manera com a base per a l'avaluació estandarditzada de futures solucions.

# Acknowledgements

I would like to express my deepest gratitude to all the people who have contributed to the completion of this thesis.

First and foremost, I would like to thank my thesis advisors, Filip Lemic and Sergi Abadal for their invaluable guidance and support throughout all the research process. His expertise and encouragement have been instrumental in shaping this work.

My sincere appreciation goes to Esteban Municio from i2CAT, and Jakob Struye from IDLab research group (University of Antwerp), who provided technical support in the development of the open source framework.

Furthermore, I am grateful to Gerard Calvo Bartra, from N3Cat for his assistance in the development of the localization machine learning model.

I would like to acknowledge the support of other key members from i2CAT, IDLab and the Data Communications and Networking group at the School of Electrical Engineering and Computer Science, TU Berlin, who kindly shared their knowledge from previous research in this area.

I would like to express my thanks to my family, friends, and colleagues for their unwavering support, understanding, and encouragement throughout my academic journey.

Last but not least, I would like to thank the staff and faculty of Universitat Politècnica de Catalunya, who have provided me with a stimulating and enriching academic environment that has fostered my intellectual growth and development.

Thank you all for your contributions, support, and encouragement.

# Revision history and approval record

| Revision | Date | Purpose |
|---|---|---|
| 0 | 09/12/2022 | Document creation |
| 1 | 10/01/2023 | Document revision |
| 2 | 20/03/2023 | Document revision |
| 3 | 15/04/2023 | Document revision |
| 4 | 25/04/2023 | Document revision |
| 5 | 01/05/2023 | Document revision |
| 6 | 09/05/2023 | Document revision |
| | | |
| | | |
| | | |

DOCUMENT DISTRIBUTION LIST

| Name | |
|---|---|
| Arnau Brosa Lopez | |
| Filip Lemic | |
| Sergi Abadal | |
| | |
| | |
| | |

| Written by: | | Reviewed and approved by: | |
|---|---|---|---|
| Date | 09/04/2023 | Date | 01/05/2023 |
| Name | Arnau Brosa Lopez | Name | Filip Lemic and Sergi Abadal |
| Position | Project Author | Position | Project Supervisor |

# 1 Introduction

During the last years, there have been many promising advancements in nanotechnology which can make a difference in the biomedical field. Specifically, nanotechnology is paving the way toward nanoscale devices with integrated sensing, computing, and data and energy storage capabilities. These devices are envisioned to be injected in the patients' bloodstreams. As such, these nanodevices will have to abide to the environmental constraints limiting their physical size to the one of the red blood cells (smaller than 5 microns). Due to such constrained sizes, their sole powering option will be to scavenge environmental energy (from heartbeats or through ultrasound-based power transfer) utilizing nanoscale energy-harvesting entities such as Zinc-Oxide (ZnO) nanowires [1]. With such devices, on the one hand, will enable the development of atomically precise materials and structures with revolutionary electrical, optical and mechanical features. On the other hand, for the first time ever, such devices will enable an interaction with living systems, bio-markers, bacteria, viruses, cancerous tissue, etc, on the same scale as they naturally interact. This opens the door to a wide range of applications, like an early diagnostics for the patients, delivering readings from the inside of the body or commanding precise drug delivery [2] [3].

However, even such interaction is not enough for enabling the envisioned applications. One of the main reasons is the lack of communication capabilities at the nano device level. Nonetheless, the emergence of graphene points to a feasible future solution. Graphene naturally resonates at terahertz (THz) frequencies (0.1-10 THz), and as such can serve as a primer for the development of THz nanoscale transceivers [4]. Although graphene-based nanoantennas have been shown to efficiently operate in the sub-THz band, this is not the only reason why this material is so important. These nanodevices need high precision and sensitivity to detect bio markers and other biological analytes. The high electron transfer rates, high charge-carrier mobility and low electrical noise levels are of utmost importance are only some of the many properties that makes graphene one of the most suited materials to use for the nano devices.

In the context of the above-discussed nanodevices, wireless communication capabilities will enable two-way communication between them and the outside world [5]. Fully integrated nanodevices with communication capabilities are paving the way toward sensing-based applications such as oxygen sensing within the bloodstream for detecting hypoxia (a biomarker for cancer diagnosis), as well as actuation-based ones such as non-invasive targeted drug delivery for cancer treatment. As recognized in recent literature, nanodevices with communication capabilities will also provide a primer for Flow Guided Localization in the bloodstream [3], [6]. Intuitively, such localization would enable associating the location of the nanodevice with a detected event, providing medical benefits along the lines of non invasiveness, early and precise diagnostics, and reduced costs [6]–[7].

Flow-guided localization is in an early research phase, with only a few works targeting the problem. The main challenges are the environment the nano devices are supposed to be deployed in. The bloodstream is a more hostile and complex environment, so the transmission is more difficult than in air or vacuum adding up the high mobility of the nanodevices within the bloodstream, with their speeds reaching 20 cm/sec, also the cen-

timeter level range of THz-based in-body wireless communication at nanoscale and energy related constraints stemming from energy harvesting as the sole powering option of the nanodevices.

This area of study has become an interesting area to research and exploit. This has led to numerous flow-guided localization proposals. Most of these studies have been done by evaluating different performance metrics and different scenarios but forgetting certain environmental variables essential to recreate a simulation of these characteristics, such as the energy consumption of the nano devices, their constraints in energy storage, self interference within the system and the constrained communication range of THz in body communication. Although all those proposals have made an encouraging progress in addressing the above challenges, some more research and further advances on such localization are needed and still to flourish.

Based on the above argument, as well as the knowledge generated through decades of research on this topic, we are convinced that there is a necessity, at least at this early stage, to design and develop a framework for objective performance evaluation of flow guided THz-based nanoscale localization. The actual research was lacking the possibility of comparing the performance of different approaches in an objective way. In other words, the reported performance results were often incomplete (e.g., targeting a single metric such as localization accuracy and ignoring the other important ones such as the latency in reporting location estimates), utilizing different performance indicators (as mean vs. median accuracy), and utilizing different evaluation criteria.

With this thesis, we aimed to move forward this area of research by proposing a framework for standardized performance evaluation of such localization approaches. Specifically, we discuss the fundamentals of flow-guided nanoscale localization, provide the categorization of existing approaches, and discuss the limitations of their current performance assessments. This is followed by proposing a workflow for standardized and objective performance assessment of flow-guided localization. In addition, an open-source network simulator is provided that implements the discussed workflow and provides the community with the first tool for realistic and objective assessment of flow-guided localization.

Our achievements include the successful development of a highly accurate simulator replicating the human body's internal environment and establishing stable two-way communication between nanonodes and anchors. This breakthrough enables event-based communications, transcending constraints and advancing healthcare technologies. Additionally, we have implemented a standardized evaluation methodology to validate the simulator's performance and gain insights into system operation, laying the foundation for future enhancements.

However, challenges persist, including relatively low accuracy due to unreliable THz communication and intermittent nanonode operation. Some regions, especially in the upper body, have similar circulation times, making prediction and training difficult. Distinguishing between paired regions poses challenges. Scalability and data disparity further impact accuracy and reliability, demanding our attention and improvement efforts.

## 1.1 Thesis Structure

The content of this thesis has been organized into several chapters to explain the work done as clearly as possible. Chapter 2 provides a brief overview of previous research and elements used in this project. In chapter 3, we dive into the technical depth of the work, starting with high-level explanations accompanied by fundamental knowledge of the project area for a better understanding of chapter 4, which focus more on the software development. Chapters 5 and 6 covers the performance assessment of the flow guided localization solution, including the definition of parameters, scenarios, and the results themselves. Finally, this thesis concludes with a discussion of the conclusions drawn from the project, its limitations, and suggestions for future work to overcome these limitations.

# 2 State of the Art of the Technology Used or Applied in this Thesis

## 2.1 Related Work

### 2.1.1 Performance Evaluation of THz Nanoscale Systems

As argued in [8], simulating the performance of a given system allows for completely controllable experimental conditions and environments. In combination with repeatability and cost-efficiency, these advantages make simulations a valuable tool to evaluate new algorithms, especially at early research stages. Given that the research on flow-guided localization is still in a preliminary stage, simulating the operation of such systems can be considered as a natural first step in the assessment of their performance. This was only meagerly recognized in the scientific community, with BloodVoyagerS [8] being the first tool for modelling the working environment of the nanodevices. BloodVoyagerS features a simplified model of the bloodstream to simulate the mobility of the nanodevices within it. As depicted in Figure 2, the simulator covers 94 vessels and organs, with the origins of the used coordinate system placed in the center of the heart. The spatial depth of all organs in the simulator is equated, with the reference thickness of 4 cm mimicking the depth of a kidney, resulting in the z–coordinates of the nanodevices being in the range between 2 and -2 cm. TeraSim [9] is the first simulation platform for modeling THz communication networks which captures the capabilities of nanodevices and peculiarities of THz propagation. TeraSim [9] is built as an ns-3 module, implementing physical and link layer solutions tailored to nanoscale THz communications. In section 4 we explain in more detail about this two open sources software.

### 2.1.2 Evaluation Methodologies for Flow-guided Localization

As argued, research lessons on the performance evaluation of indoor localization systems can to an extent be applied for objective and standardized assessment of flow-guided localization. The EU EVARILOS project was among the early efforts aiming at such performance assessment for RF-based indoor localization [10]. Within the project, a performance assessment methodology was developed, which included a number of evaluation scenarios, envisioned capturing the performance of evaluated solutions along a heterogeneous set of metrics including localization accuracy, latency, and energy consumption, and assessing and mitigating the negative effects of RF interference on the performance of the evaluated solutions. The project also yielded a web platform populated with raw data that can be inputted in an indoor localization solution for its streamlined performance assessment along a number of standardized scenarios. Similar approach was followed in the NIST PerfLoc project [11], however with a set of possible solutions to be evaluated extending beyond only Radio Frequency (RF) to Inertial Measurement Unit (IMU)-based, Global Positioning System (GPS)-supported, and other hybrid approaches. Finally, the IPSN/MICROSOFT Indoor Localization Competition [12] was among the first efforts to support back-to-back evaluation of different indoor localization approaches along the same set of conditions.

The above-discussed and consequent efforts yielded the following lessons: i) performance

comparison of different indoor localization approaches can be carried out in an objective way by following the same evaluation methodology, i.e., utilizing the same environments, scenarios, and evaluation metrics, ii) such evaluation can be streamlined by providing a set of raw data captured along a standardized evaluation methodology, which is envisioned to be used as an input to an indoor localization solution, and iii) the performance of RF-based indoor localization can be degraded by both self-interference and interference from neighboring RF-based systems operating in the same frequency band.

In the current outlook on the performance assessment of existing flow-guided localization, the approaches from [7] and [13] are evaluated in a rather simplified way accounting solely for the mobility of the nanodevices as modeled by the BloodVoyagerS. As such, their performance assessments ignore many potential effects of wireless communication (e.g., RF interference), as well as energy-related constraints stemming from energy-harvesting and, consequently, the intermittent operation of a nanodevice [1]. It is also worth mentioning that [6] carried out a limited performance evaluation assessing the number of nanodevices needed for localizing a nanodevice at any location in the body in a multi-hop fashion. The derived assessments can, therefore, at this point only serve as a rough indication due to their low levels of realism and subjective evaluation methodologies. In this work, we enhance the realism of such assessments by jointly accounting for the mobility of the nanodevices, in-body nanoscale THz communication between the nanodevices and the outside world, and energy related and other technological constraints (e.g., pulse-based modulation) of the nanodevices.

# 3 High Level Design

In this section, we will cover non-technical aspects of our localization solution, providing a high-level explanation of the flow-guided localization approach. First, we will briefly introduce the fundamentals of flow-guided localization, followed by an overview of the proposed solution.

## 3.1 Flow Guided Localization Fundamentals

The different localization solutions can be categorized according to the type of application they support. Obviously, there is a need to be able to locate some nano devices that are mobile. These devices target three different types of applications [14]: First is the localization of macro scale elements inside the body, specifically for localizing gastric capsules (as there is a clear diagnostic benefit of assigning the measurements of the gastrointestinal system with the locations at which they were taken) and implants (for detecting their movements away from the intended deployment locations). It is expected that these devices are not intended to be of nanoscale magnitude and have reduced mobility (few cm/h in the gastrointestinal system) or simply no mobility in the case of implants. This reduces the complexity of locating the devices in contrast to the following two categories. The second category includes the localization of nanoscale devices that have reduced mobility levels and are used in applications such as tracking fiducial markers (devices that provide accurate target location for tumors or organs which move in respect to surrounding anatomy) and other types of miniaturized implants. Although many of these applications do not specifically require their dimensions to be nanoscale, in order to enable all its applications these devices must be nanoscale. (serve as example an early targeted treatment of small-scale tumors). Here is a representation of this category [6] which consists of nanodevices deployed and flowing through the bloodstream. In this scenario an interactive localization is proposed in which the nanodevices closest to the surface are located first thanks to anchors that will be on the surface of the body, using the nanodevices located first to locate the ones deeper in the body. Such an approach could conceptually be applied for localizing nanoscale implants within the body.

The final category is the flow-guided nanoscale localization considered in this work. Here, the goal is to use the nanodevices to detect and localize a target event, not necessarily to localize themselves. The work in [6] can conceptually support this type of scenarios, therefore, that is why it is included in both categories in figure 1. Even so, the works that are the models for this category are [13] and [7]. In these two works, machine learning algorithms are used to distinguish the region of the body through which a nanodevice passes during one circulation through the bloodstream. In [8] base this procedure on tracking the distances traversed by a nanodevice in its circulations through the bloodstream by utilizing a conceptual nanoscale IMU. This process therefore generates a series of problems, such as the storage for processing the IMU-generated data, also must be taken into account the sharp turns of the nanodevices in the bloodstream that will adversely affect the IMU measurements. In [7] an attempt is made to solve these problems by tracing the time it takes to make a circulation and then sending this data to an anchor close to the heart using THz-based back-scattering at the nano device level.

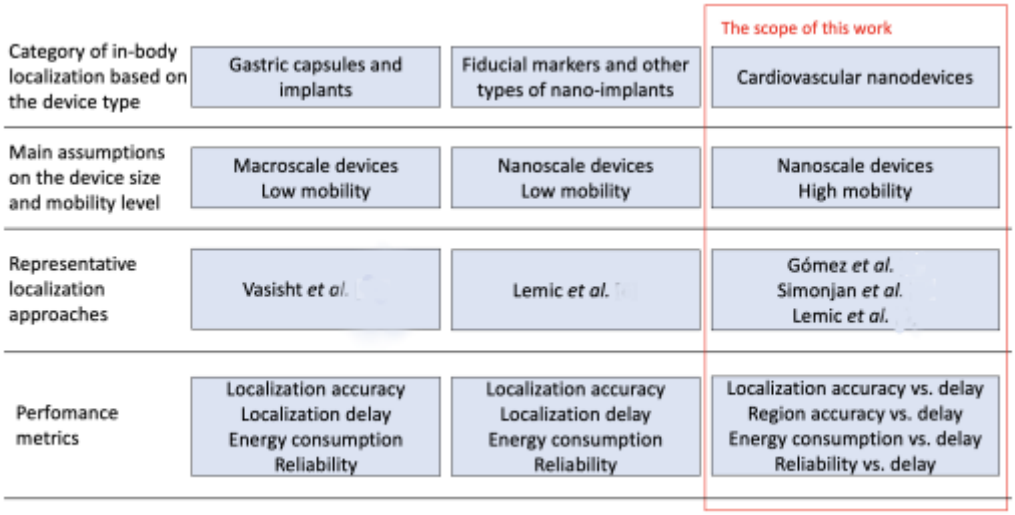



Figure 1: Scope of the present work

In these models, the region through which the nanodevice has circulated is detected and not the specific point where there is an event as in [6]. Although being able to locate the specific point would give tremendous benefits for health-care diagnostics, the accuracy and reliability in these cases is not acceptable, but with the models [13] and [7] we can obtain an increase in these performance metrics with an increase in the number of circulations made by the nanodevices in the bloodstream. On the other hand, this increase causes the energy consumption levels to rise. Therefore, in flow-guided localization the relevant performance metrics such as the point and region accuracy, reliability, and energy consumption should be considered as a function of the application-specific delay allowed for localizing target events.

## 3.2 Experimentation Overview

Our system relies on bilateral communication between two main elements: the nanonodes and the anchors. The nanonodes are small devices that flow through the bloodstream and constantly collect and send data to the anchors, which are strategically located on the skin. The anchors transmit beacon packets and receive backscattered responses from the nano nodes.

The nanonodes are equipped with capacitors for energy storage and ZnO nanowires for energy harvesting. The capacitor charging is modeled as an exponential process, taking into account the energy-harvesting rate and interval (e.g., 6 pJ per 20 ms for harvesting from ultrasound-based power transfer), which charges the capacitor more or less depending on the available energy at the moment.

Due to energy harvesting limitations, the nanodevices exhibit intermittent behavior, turning on when their energy levels are above a certain threshold and turning off when depleted. If turned on, they periodically execute sensing or actuation tasks at a given frequency, with each task consuming a constant amount of energy.

These sensing tasks are designed to detect events, such as biomarkers, which are hard-coded into the experiment. The locations of these events are varied in different scenarios to test the accuracy of the system in different parts of the body. The conditions for the nanonodes to detect events include:

1. The Euclidian distance between the actual nano node position and the event position is less than predefined threshold.
2. The nano node is turned on

The communication between an anchor and a nanodevice involves passive reception of a beacon, followed by active transmission of a response packet from the nanodevice. This means that nanodevices do not spend energy on receiving, but they do for sending data to the anchors. Although we attempted to implement full back-scattering communication, the in-body attenuations were too high, preventing the packet from reaching the anchor. To resolve this issue, we established a passive reception of the beacon sent by the anchor and an active transmission of the response. This method allows the nanonodes to spend energy on sending a packet back to the anchor and sensing events. However, this type of communication leads to a large energy drain, which is why an energy harvesting model is crucial.

The anchor transmits a constant beaconing frequency and power. Each beacon packet contains the anchor's Medium Access Control (MAC) address and ID. In the backscattered packets, the nanodevices report their MAC addresses, the time elapsed since their last passage through the heart, and an event bit. These data points represent the raw data that can be used in a flow-guided localization approach to locate a target within the body. We process this raw data in evaluations to determine the target's region and evaluate the system's performance.

Flow-guided localization of nanodevices in the bloodstream requires at least one anchor mounted on the patient's body. With a single anchor strategically positioned near the heart, flow-guided localization approaches in [13] and [7] can be enabled because nanodevices are guaranteed to pass through the heart in each iteration through the bloodstream. In this thesis, we evaluate the performance of the system with just one anchor. Although the simulator supports additional anchors by specifying their positions in the configuration file.

Each time a nanodevice passes through the heart, the time elapsed since the last passage is re-initialized to zero to prevent compounding multiple iterations. The event bit is a logical 1 if a target event is successfully detected and 0 otherwise. Similarly, the event bit is re-initialized to 0 in each passage through the heart.

# 4 Low Level Design

In this chapter, we will delve into the technical aspects of our thesis, presenting a detailed low-level design of our flow-guided localization solution. First, we will provide an overview of the software tools used to develop our solution, their interactions with each other, and the results they produce. Next, we will define all the internal and external components that make up our simulator. To conclude, we will describe the workflow of our simulator in detail.

## 4.1 Utilized Tools

In chapter 2, we discussed how our localization system integrates three different open-source software packages. The integration of these three software packages provides the research community with a powerful tool for realistic and objective assessments of flow-guided localization. This tool will facilitate the evaluation of new algorithms, particularly in the early stages of research, and will enable many advancements in future research in this area.

### 4.1.1 Blood-Voyager-s

The first step of our procedure involves utilizing the Blood-voyager-s software, which was introduced in section 2. This software provides a simplified model of the human cardiovascular system, allowing us to determine the global positions of the nanonodes within the system. We integrated this software into our NS-3 code and generated a csv file containing the positions of the nanonodes, which we will later use in our main simulation. To adapt the csv file to our specific needs, we developed a python script for processing the data. The resulting cardiovascular positioning map, used for simulating the global positions of the nanonodes, is displayed below.

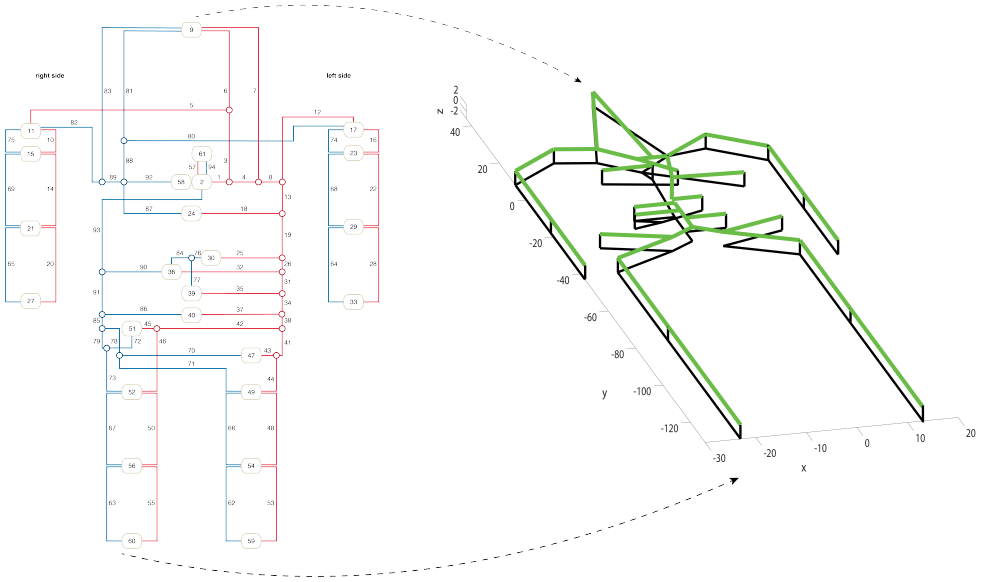



Figure 2: Blood-voyager-S positioning structure

Figure 2 shows the main vessels that provide local coverage and contain large organs and limbs. Overall there are 94 vessels and organs covered. The human body is three-dimensional, the spatial direction from the front (anterior) to the back (posterior), the z-coordinate, is still lacking. The measurement of the organs from anterior to posterior on the skin is not possible and again there is no record of all relevant dimensions. Therefore, the spacial depth is established using as a guideline, the thickness of the kidney, which is between 3 and 5 cm, then in this simulations, all z distances are set at 4 cm, as explained in chapter 2.

The simulator further assumes that the arteries and veins are set anterior and posterior, respectively. Transitions from the arteries to veins happen in the organs, limbs, and head. In the heart, the blood transitions from the veins to arteries. The flow rate is modeled through the relationship between pressure difference and flow resistance. This results in the average blood speeds of 20, 10, and 2–4 cm/sec in aorta, arteries, and veins, respectively. Transitions between the arteries and veins are simplified by utilizing the constant velocity of 1 cm/sec.

To summarize this section, we have used Blood Voyager-S to obtain the global position of the nanonodes every second throughout the simulation. By doing this, we will be able to integrate these positions into our nanonodes in the simulator.

### 4.1.2 NS-3+TeraSim

The main part of our develop solution has been made using NS-3 and TeraSim. NS-3 is an open source computer networks simulation environment that is based on discrete-event simulation. In such a simulator, each event is associated with its execution time, and the simulation proceeds by executing events in the temporal order of simulation time. It is designed for research use, thinking of the needs of the research community, and fosters a community-based collaboration model both to support the development of new modules and to perform validation or peer review activities across different groups. NS-3 is distributed under an open-source model and a free software paradigm.

TeraSim is another open source code, compatible with NS-3, and it's the first simulation platform for THz communication networks which captures the capabilities of THz devices and supports all the peculiarities of the THz communications, as the THz channel or the molecular loss absorption among other features [9]. Specifically, at the physical layer the simulator features pulsebased communications with an omnidirectional antenna over distances shorter than 1 m, assuming a single, almost 10 THz wide transmission window. At the link layer, TeraSim implements two well-known protocols, ALOHA and CSMA, while a common THz channel module implements a frequency selective channel model, assuming in-air wireless communication. We will utilize BloodVoyagerS and TeraSim as the starting point in the development of the envisioned simulator.

With these two open source softwares we have managed to design a communication system between the nanonodes and the anchors simulating the environment inside the human body. This solution will integrate all the features and constraints discussed in chapter 1 and 3 necessary for both the nanonodes and the anchors to operate correctly within such a conflicting environment as the in-body is.

## 4.2 System Design

In this section, we will provide a detailed explanation of the internal structure of our nanonodes in our simulator, which consists of various modules. We will delve into the main features and functionalities of each module. Additionally, we will describe the different external modules used to establish bilateral communication between the nanonodes and anchors.

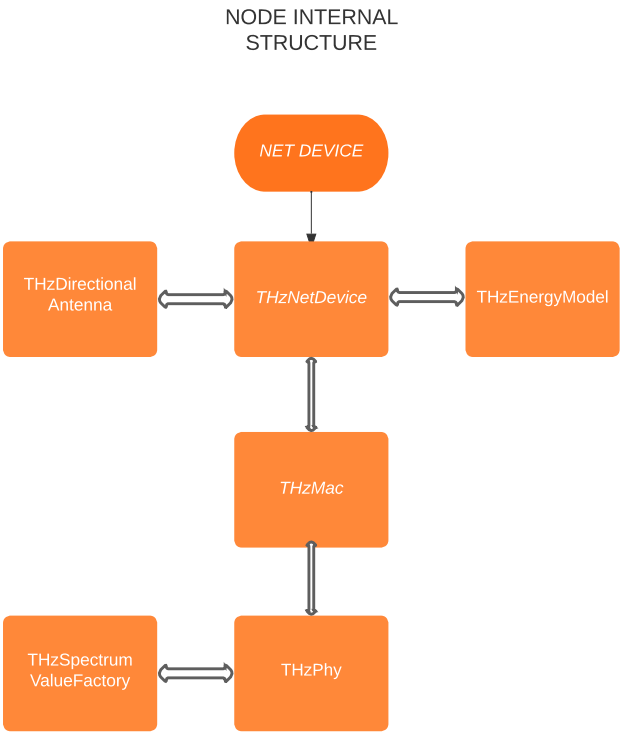



Figure 3: Nanonode structure

Figure 3 depicts the internal structure of both the nanonodes and the anchors. The design includes a physical and link layer, as well as several assisting modules such as an energy model that addresses the constraints discussed in previous chapters. Additionally, there are other modules that we will now describe in greater detail.

1. **THzNetDevice:** THzNetDevice is derived from the base class NetDevice provided by ns-3 to create new MAC protocols. It performs as the joint point which connects the THzChannel module, THzPhy module, THzMac module and the assistant modules such as the THzDirectionalAntenna and THzEnergyModel together.
2. **THzMAC:** At the link layer there is a MAC protocol implemented. When a packet is enqueued in the THzMacNano, if the node is turned on then it passes the packet over the physichal layer for 0-way (ALOHA) handshake protocol. A node receives

the packet then. The desired receiver then sends an acknowledgement (ACK) packet to the sender.

3. **THzPhy:** The physical layer for the nanoscale transmit by using TSOOK, a modulation scheme based on the transmission of one-hundred-femtosecond-long pulses by following a non-off keying modulation spread in time.
4. **THzDirectionalAntenna:** The THzDirectionalAntenna module based on CosineAntennaModule, which is provided in the NS-3 platform. The antenna gain of the cosine model We as we are doing simulations at a nanoscale level so we are using an omnidirectional antenna.
5. **THzEnergyModel:** The nanonodes use an energy model, unlike the anchors which are assumed to have enough energy for transmitting. This energy model uses passive reception, which means that there is no energy consumption for receiving a packet. However, there is energy consumption for sending back the packets to the anchors and for performing sensing tasks. To manage energy usage, we have implemented an ON/OFF threshold. The nano nodes begin the simulations with a maximum storage capacity of 800 pJ. When the energy of a nano node is depleted to 0 pJ, the node is put into an OFF mode, where it is unable to send or perform sensing tasks. The nanonodes will recharge over time, and once they reach or exceed the ON threshold, which we have defined as 10 pJ, the nanonode will return to the ON state.

   Regarding our energy harvesting model, the nanonode will harvest energy exponentially at a constant time. Our energy harvesting model is designed to be highly efficient. The nanonodes operate on an exponential harvesting curve, meaning that the amount of energy harvested will increase rapidly when their actual energy is at the lowest levels and increase slower as their energy storage is being fulfilled. Additionally, this harvesting process occurs at a constant rate, allowing for reliable and consistent energy generation.

   The harvested energy can be specified with the duration of the harvesting cycle $tcycle$ and the harvested charge per cycle $\triangle\ Q$. Capacitor charging through energy harvesting can be accurately modeled as an exponential process, accounting for the total capacitance $C_c$ of the nanonode, i.e., $C_c$ depends on the maximum energy storage capacity $Emax$ and the generator voltage $V_g$ . In the modeling, it is required to know in which harvesting cycle $ncycle$ the nanonode is, given its current energy level $Encycle$

$$n_{cycle} = (V_g * C_c)/\ \triangle\ Q * ln(1 - \sqrt{2E_{ncycle}/C_c * V_g^2}) \tag{1}$$

$$E_{ncycle+1} = (C_c * V_g^2)/2 * (1 - \exp -(\triangle\ Q * (n_{cycle+1}))/V_g * C_c)^2 \tag{2}$$

Where $C_c$ is:

$$C_c = 2 * E_{max}/Vg^2 \tag{3}$$

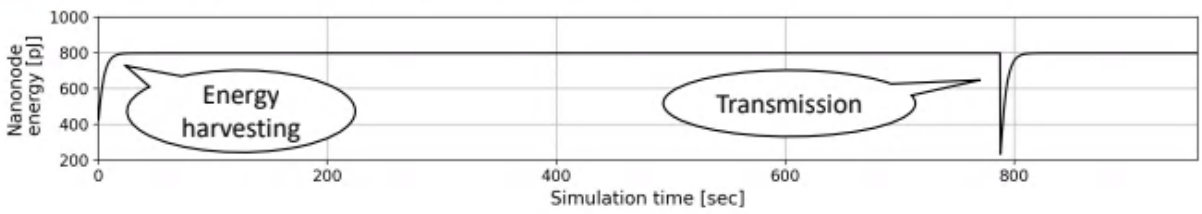



Figure 4: Energy model simulation

Figure 4 provides a detailed overview of the behavior of our energy model. Specifically, the nanonode operates on a constant harvesting interval of 20ms, leveraging the exponential formula described earlier. The figure demonstrates that the nano node is capable of harvesting more energy when the available energy is lower, and that it continues to harvest energy until it reaches its storage capacity limit. Once the nano node sends a packet back to the anchor, its energy level begins to decrease, as illustrated in the figure.

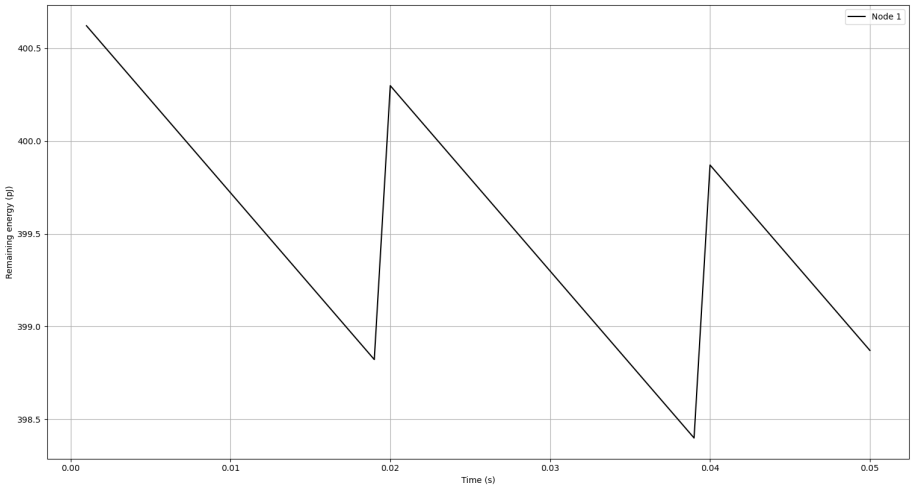



Figure 5: Event detection consumption

Figure 5 above illustrates the energy consumption required for event detection. To demonstrate this performance, we employed a granularity of 1000 samples per second. Specifically, we utilized an energy consumption rate of 0.1pJ per event detection. As depicted in the figure, we observe that the nanonode consumes 0.1pJ for each sensing task.

Table 1 shows the values predefined in our simulations for the following parameters defined in the formulas above.

| Variable | Value |
|---|---|
| Harvesting cycle duration | 20ms |
| Harvested charge per cycle ($\triangle Q$) | 6pC |
| Generator voltage($V_g$) | 0.42V |
| Turn OFF threshold | 0pJ |
| Turn ON threshold | 10pJ |
| Maximum energy storage capacity (Emax) | 800 pJ |

Table 1: Energy model parameters

Having outlined the modules at the nanonode level, we will now provide an overview of the external modules that enable bilateral communication between the nanonodes and anchors.

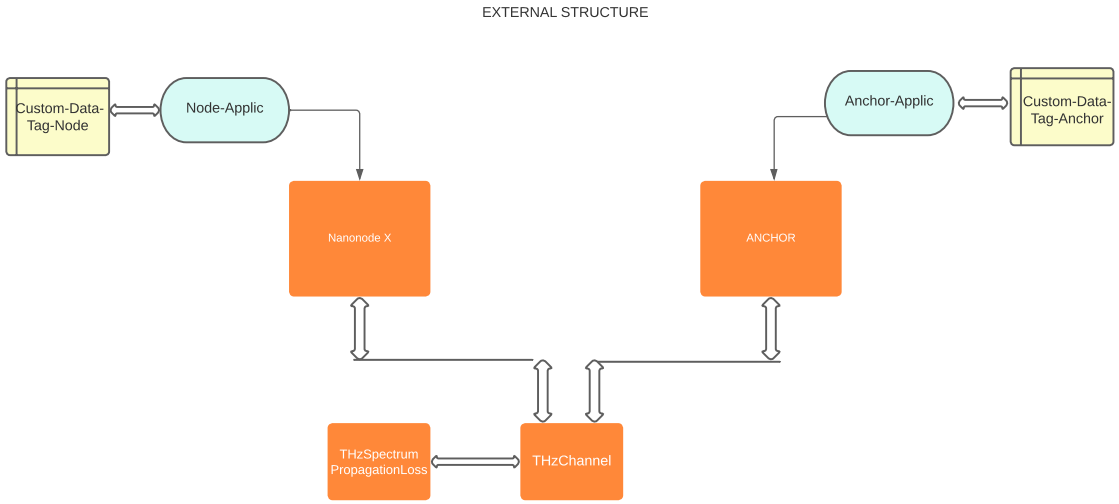



Figure 6: Simulator structure

Figure 6 illustrates the various modules that interact with both the nanonodes and anchors to enable effective communication between both.

1. **Node-Applic:** This module acts as the application and transport layer of the nanonodes, because it manages the behavior of the nanonodes and also generates the packets that will be sent to the anchors. The main functionalities of this module are:
   (a) Receiving the beacons from the anchors and substract their addresses and IDs.
   (b) Checking it there is enough energy for continue sensing new events.

(c) Estimating the euclidean distance between the actual nanonode position and the event position which is hardcoded. The threshold that determines whether the event will be detected or not will vary depending on the scenarios but we will establish as a base value that if the Euclidean distance is equal to or less than one, then the event will be detected. We used the following equation to estimate the Euclidean distance:

$$d(p,q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2 + (p_z - q_z)^2} \tag{4}$$

p,q being the actual position of the BNS and the event position respectively.

(d) Sending the response back to the beacons with all the raw data necessary for later use that data to locate in which region of the body the target was, and obtain the performance of our system.

2. **Anchor-Applic:** It has the same functionality as the one explained before, but this is applied for the anchors. It manages the behavior of the anchors and also generates and prepares the packets that will be sent to the nanonodes. The anchors will generate beacons every 100ms and send it to their surroundings waiting for nanonodes to enter its communication range and receive the packets. This beacons are sending the anchors' ID and their addresses, so that the nanonodes can know where they have to send back the response. It is also the responsible to receive the response packets of the nanonodes with all the raw data.

3. **THzChannel:** The THz channel takes into account waveforms with realistic bandwidth and applies the frequency selective propagation behavior of the THz channel. When the waveform generated by THzSpectrumValueFactory is passed down to the channel, THzChannel first obtains the antenna gains of the transmitter and receiver from the THzDirectionalAntenna module Then it passes the waveform, total antenna gain and the mobility model to THzSpectrumPropagationLoss and finally passes the packet along with the received power to the physical layers of the receiver to simulate the wireless broadcast channel.

4. **THzSpectrumPropagationLoss:** This module together with the Inbody-loss module, model and recreate the attenuation of the environment of in body. The propagation at THz band is mainly affected by molecular absorption which results in both molecular absorption loss and molecular absorption noise. We calculated the path loss of the vessels, tissues and skin, based on certain parameters that we are explaining now. The total path loss for an in-body transmission/reception is:

$$TotalPathLoss = VesselPathLoss + TissuePathLoss + SkinPathLoss \tag{5}$$

To calculate each path loss for each part is:

$$PathLoss = SpreadingLoss * AbsorptionLoss \tag{6}$$

The necessary variables for getting those path loss are:

(a) The carrier frequency which we established in our simulations as 10THz

(b) The distance between the anchor and the nanonode

(c) Different constants like the thickness of the skin, vessels, tissues or the speed of light

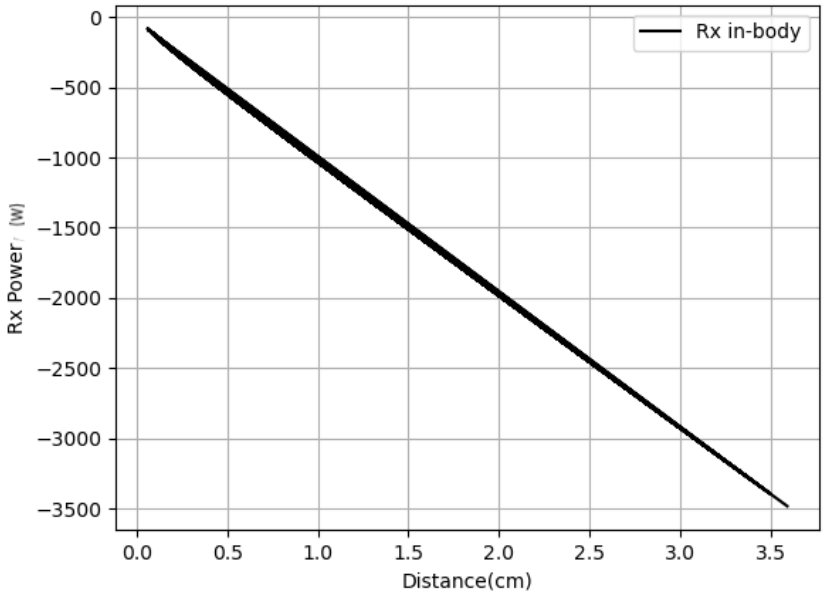



Figure 7: Nanonode reception power

Figure 7 shows the performance of the receive power according to the distance between the nanonode and the anchor. It is reflected how the received power decrease a lot in only a few centimeters due to all the attenuation the in-body communication have. The communication range is around 0.4/0.5 cm.

5. **Tag Classes:** The both tag classes (Custom-Data-Tag-Node and Custom-Data-Tag-anchor) encapsulates the information that we want to sent in NS-3 objects called Tags and that will be used by the application layer to generate the packets. Custom-Data-Tag-Node encapsulates the raw data needed to send to the anchor and Custom-Data-Tag-anchor encapsulates the data needed to send to the nanonodes, the receivers in the application level are the ones in charge of untag that packets.

## 4.3 Workflow

The architecture of the simulator follows a well-established ns-3 layered model, as depicted in Figure 8. The Anchor-Application module implements continuous beaconing with a predefined period (nb., with 100 ms being a default value) to broadcast with their address and ID. Each beacon packet is forwarded to the THzNetDevice module. At the THzNetDevice the packet calls the enqueue method from the THzMac, in the MAC protocol a header is included, after that it passes the packet to the physical layer, THzPhy

creates its own data structure for the packet and adds it to the list of ongoing transmissions. If a new transmission or reception comes while a transmission is going on, THzPhy will use the information to interleave and check for collision. The link and physical layers implement the ALOHA protocol and TS-OOK modulation, respectively.

Then the packet is passed to the THzChannel by calling the SendPacket() method of the channel through the handle (pointer to THzChannel).The THz channel is modeled by calculating the receive power for each communicating pair of devices and scheduling the invocation of the ReceivePacket() method accounting for the corresponding propagation time. The channel model entails in-body path-loss and Doppler term. The path loss is calculated using the attenuation and thickness parameters of the vessel, tissue, and skin. The Doppler term is accounted for by evaluating the change in relative positions between the nanodevices and anchors with time.The ReceivePacket() method of the channel invokes the RecivePacket() method of the THzPhy using the pointer to it.

The ReceivePacket() method of the THzPhy checks for the collision and calculates the SINR every time there is a new interfering signal. Then it marks the packet as collided if the SINR value is less than the predefined threshold for detection, otherwise it marks the packet as not collided. The ReceivePacketDone() of the PHY then checks the completely received packet to see if it was collided i.e., the SINR was less than the threshold at some point during reception. If it did collide, the packet is dropped. Otherwise, the packet is passed to the MAC layer. The MAC layer first checks if it is the intended receiver of the packet by looking into the destination address. If it is the intended receiver and the packet is received correctly it passes the packet to the upper layer. Finally at the application layer the receiver (anchor) untag all the information that was at the packet. At the nanodevice level, the receive power of the beacon is used for setting up the transmission power of the packet to be backscattered. This is followed by backscattering the response packet from the nanodevice toward the anchor by utilizing the same procedure as for the transmission of the beacon.

The anchors are assumed to be static entities and feature sufficient energy for continuous operation. The nanodevices are assumed to be energy-harvesting entities that are mobile within the bloodstream. To model their mobility, we have integrated BloodVoyagerS in our simulator. Invoking a BloodVoyagerS execution results in generating csv file that specifies the locations of the nanodevices in the bloodstream. The positions generated in these simulations are generated with a frequency of 1 Hz. Since ns-3 is an event-driven simulator, at each BloodVoyagerS-originating location of a nanodevice, the nanodevice is assumed to carry out a sensing/actuation task. Given that for certain applications carrying out such tasks could be required more frequently, we provide an upsampler for BloodVoyagerS-originating locations sampled at 1 Hz. As the vessels in BloodVoyagerS are modeled using straight lines, the upsampling is based on linear interpolation with a small random component drawn from a zero-mean Gaussian distribution, representing vortex flow of blood and minor changes in the diameters of veins, arteries, etc. At each new location, the nanodevice is expected to carry out a task for detecting an event of interest.

Finally, once the communication system between the nodes and the anchors has been

established, and they are able to exchange messages correctly, then we need to converting the raw data that the nano nodes gathered and sent to the anchors into a usable solution for our flow guided localization solution. This raw data is, the time elapsed since their last passage through the heart, and an event bit.

For each scenario proposed, this raw data is generated. Once we have all the data necessary, then we need to evaluate the performance of our simulator. For that we are going to use an ML algorithm. Finally, after applying the ML solution, we can evaluate the performance of the different scenarios, and have an initial idea of the limitations of our simulations and necessary future work.

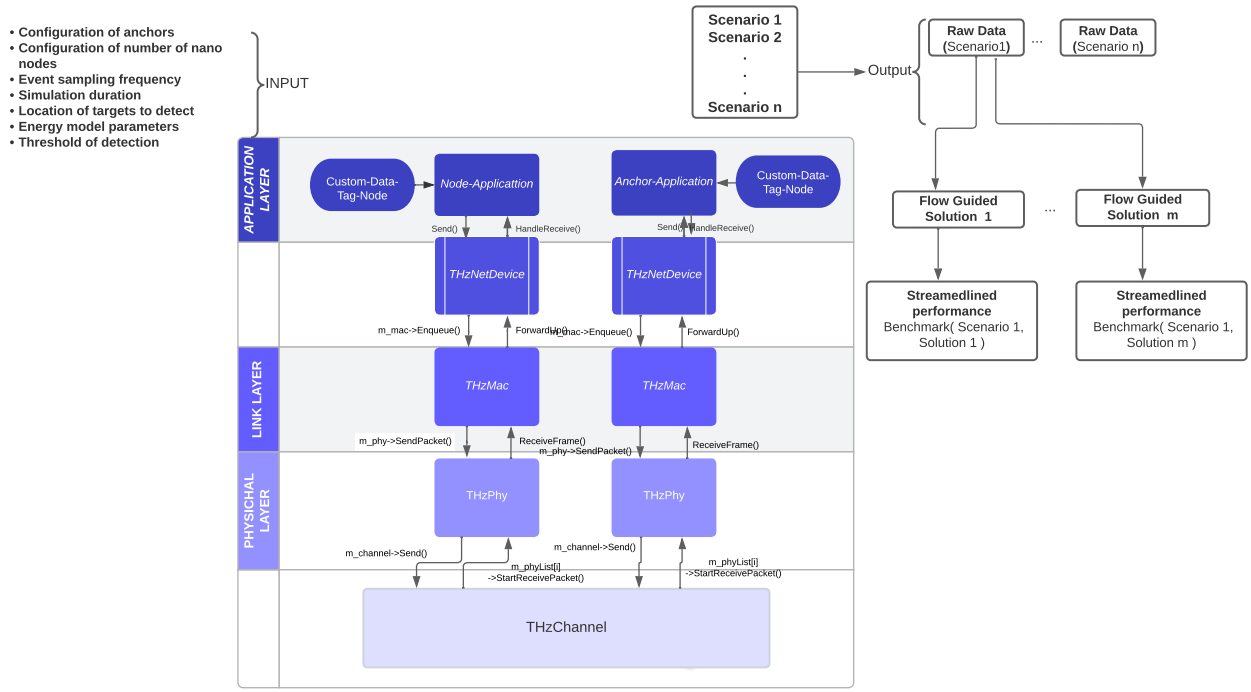



Figure 8: Simulator workflow

Figure 8 presents a summary diagram of the workflow described in this section. The input signals consist of the parameters we have designed, which can be readily modified in the configuration files. The bilateral communication between the anchors and the nano nodes is initiated at the application layer of each node/anchor and continues down to the corresponding application layer of the other node/anchor. The resulting raw data is then utilized to generate our flow guided solution.

# 5 Flow Guided Localization Solution

Once communication between the nanodevices and the anchors has been enabled, and the nanodevices are able to send the necessary data to the anchors (such as circulation time, nanodevice ID, and whether the event has been detected or not), it is necessary to process the large amount of data that is generated. For this purpose, we have created a set of Python scripts that will be responsible for processing all this data. Once we have the processed data, it will go through an ML system that we have designed, which will allow us to obtain a prediction of the regions where the events are located and finally, we will be able to obtain an evaluation of these results.

## 5.1 Flow-guided Localization Solution under Test

As we have mentioned before, in order to achieve our flow guided localization solution, we need to run the data collected during our simulator through a machine learning algorithm that we have designed. This algorithm will predict the region where the events are located based on their circulation times (which are transmitted by the nano nodes as they pass through the heart's anchor).

A large number of simulations had to be carried out in order to train this algorithm. Our flow guided localization solution currently distinguishes 25 regions of the human body, defined in table 2. To train the model, various training simulations were conducted in which the position of the event to be detected was varied. This enabled us to obtain a large amount of data with circulation times for each region.

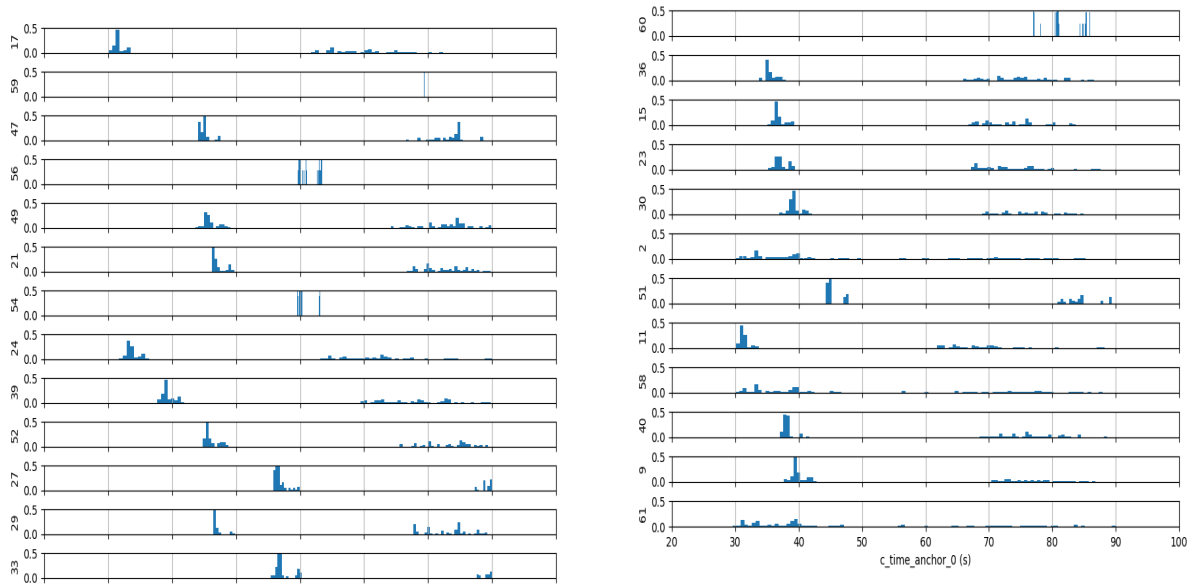



Figure 9: Distribution of circulation time on anchor 0 for each region

In Figure 9, we can observe the circulation times of each of the regions that we have collected through the various simulations we have conducted to train the model. The

different regions are defined by an ID. In table 2, we associate each ID with the respective name of the region. One problem encountered in training this model has been the disparity of data obtained from the different regions, which made it difficult to have a functionally trained model. To address this, we had to conduct many more simulations of the regions from which we obtained less data to achieve a balanced number of data in all regions.

| Regions ID | | | |
|---|---|---|---|
| ID | Region | ID | Region |
| 17 | left shoulder | 33 | left hand |
| 59 | left foot | 60 | right foot |
| 47 | left pelvis | 36 | liver |
| 56 | right knee | 15 | right upper arm |
| 49 | left hip | 23 | left upper arm |
| 21 | right elbow | 30 | spleen |
| 54 | left knee | 2 | left heart |
| 24 | thorax | 51 | right pelvis |
| 39 | intestine | 11 | right shoulder |
| 52 | right hip | 58 | right heart |
| 27 | right hand | 40 | kidneys |
| 29 | left elbow | 9 | head |
| 61 | lungs | | |

Table 2: Regions ID

In Figure 9, we could already observe how some data points in several regions had circulation times that were too high and far from the average. In Figure 10, we extracted the circulation times of a single nanodevice in a single simulation. The main takeaway from both figures is that, for some raw data instances, the circulation time is larger than 90 sec, which is the maximum circulation time that might occur in a single loop trough the bloodstream. This implies that in some circulations the raw data is not reported to the anchor and, when the data is eventually reported, it contains the compound of multiple such circulations. Such behavior is a result of one of the following:

1. Intermittent operation of a nanonode due to energy-harvesting, resulting in the nanonode sometimes not featuring sufficient energy for sensing or transmission
2. Self-interference from the other nanonodes and anchors, resulting in reception and transmission errors.

In addition, random paths of the nanonodes in the vicinity of the target event (i.e., in an organ, limb, or head) can result in the nanonodes missing the event due to its Euclidean distance from the event never being smaller than the threshold of 1 cm, despite the fact that they went through the loop that contained the event. This implies that the event bit parameter might in some cases be erroneous.

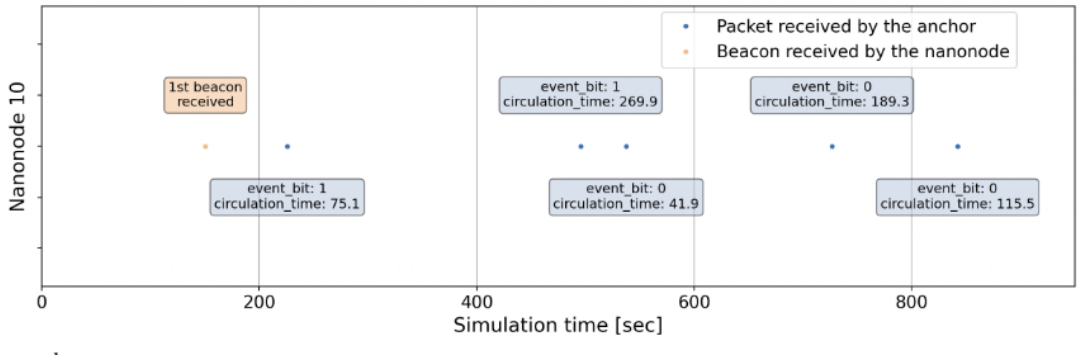



Figure 10: Raw data output

## 5.2 Design Space

We are conducting experiments with various scenarios based on specific parameters that are relevant to our research. By doing so, we aim to obtain diverse performance metrics and outputs, which will help us draw better conclusions about the strengths and weaknesses of our localization solution. Furthermore, this will inform us about how we can improve our model for future studies. All of the parameters we designed can be modified and set in the "executeProgram.sh" file, which serves as the foundation for running the system explained in section 4. The parameters include the number of nano nodes deployed in the cardiovascular system, the number of anchors used as receivers, as well as their optimal placement on the body, the sample granularity, and the detection threshold for events. In the following sections, we will provide more detailed information on each of these parameters.

1. **Number of nano nodes:** The number of nano nodes deployed in the cardiovascular system complicates the simulation run time as the amount of data to be handled increases. However as more data we should to analyse more accuracy we will have on the results.The number of nanonodes deployed in the different scenarios will have a range of values according to the number of full bits. By this we mean that the numbers of BNS will be of the order of $2^X$ where X is an integer. The reason why we implement these values is because in this way we can optimize to the maximum the number of bits destined to designate the ID of the nano nodes. We have also calculated the simulation run time for the various scenarios in order to quantify the amount of resources it consumes computationally.

2. **Granularity** We have explained that we obtain the global positions of the nanonodes through simulations made with the blood-voyager-S software. These simulations provide us with position updates every second until the final simulation time. But in elements as mobiles as the nanonodes are, getting updates every second may not be enough, so we believe, it may be interesting to add this granularity variable. With this variable what we have done is to interpolate positions according to the chosen option.To interpolate these positions, we have used a linear interpolation method added to a small random component as well (representing vortex flow of

the blood, arbitrary changes in the diameter of the veins and arteries, etc.), to represent that we added some random numbers from a zero-mean Gaussian distribution. Below you are going to have the interpolation formulas used.

$$coord_x = [..., x_1, x_{i1}, x_{i2}, x_2, ...] \tag{7}$$

$$coord_y = [..., y_1, y_{i1}, y_{i2}, y_2, ...] \tag{8}$$

$$coord_z = [..., z_1, z_{i1}, z_{i2}, z_2, ...] \tag{9}$$

Being $t_i, x_i, y_i, z_i$ the interpolated data and

$$x_{i1} = x_1 + A_{int}|x_2 - x_1|/(N_{int} + 1) \ + Gaussian(0, |t_1 - t_{i1}|/10) \tag{10}$$

$$x_{i2} = x_1 + A_{int}|x_2 - x_1|/(N_{int} + 1) + Gaussian(0, |t_1 - t_{i1}|/10) \tag{11}$$

Being $N_{int}$ the number of interpolated rows and Aint the actual interpolation row that goes from 1 to $(N_{int} - 1)$ and that's applied for every axe and every position that we have to add according the option we choose to analyse in each scenario.

3. **Event threshold detection** The detection of an event is done by calculating the euclidean distance between the actual position of the nano node and the position where the event is located. The event threshold detection is the maximum distance that the nano nodes can detect the events. Obviously, the number of events detected will increase proportionally to the detection range. However, the constraints due to the size of nano nodes and their sensors limit that detection range. We have done several tests changing that threshold in order to evaluate its performance.

## 5.3 Parametrization of Scenarios

As we mentioned in chapter 5, we have automated several parameters so that we can conduct various simulations, changing the variables and thus obtain performance based on the variables we have discussed in this chapter. The parameters that have been automated are: the number of nanonodes, the number of anchors, granularity, simulation time, the number of events per region and we also are able to modify their location using some functions we designed.

To perform the training simulations of the machine learning model, we have defined the different parameters, explained above, as follows:

| | |
|---|---|
| number of nanonodes | 64 |
| number of anchors | 1 |
| granularity | 1/3s |
| simulation time | 2000s |

Table 3: Hyperparameters

We designed the evaluation in a way that we can understand how the parameters affect the performance metrics. The first evaluation consists in using a baseline scenario, which contains, the base values (chapter 5.3.1) we designed to evaluate a first set of results and analyse its behaviour. Once we have the baseline scenario, we divide our evaluations in three scenarios:

1. Based on the number of nanonodes: In this evaluation we modify the number nanonodes deployed on the bloodstream. The range of values to test in this scenario are: 32,128,250 nanonodes keeping the other parameters as the baseline scenario. A total of 4 scenarios in this set.
2. Based on the granularity: In this evaluation we modify upsampling frequency. The range of values to test in this scenario are: 1 sample/s, 2 samples/s, 5 samples/s and 10 samples/s nanonodes keeping the other parameters as the baseline scenario. A total of 4 scenarios in this set.
3. Based on the event threshold: In this evaluation we modify the threshold of detection. The range of values to test in this scenario are: 0.5cm, 2cm and 3cm keeping the other parameters as the baseline scenario. A total of 3 scenarios in this set.

We have conducted 25 simulations for each scenario of a set, each of them moving the event to be detected to a different region. As we have a total of 25 regions, this means 25 different simulations.

In table 4 we have the range of values we evaluate in each set, including the baseline values.

| **Number of nano nodes** | **Number of anchors** | **Granularity** | **Event threshold detection** |
|---|---|---|---|
| 32 | 1 at the heart | 1 samples/s | 0.5cm |
| 64 | | 2 samples/s | 1cm |
| 128 | | 3 samples/s | 2cm |
| 256 | | 5 samples/s | 3cm |
| 512 | | 10 samples/s | |

Table 4: Parameters range

For our first scenario we have decided to choose a series of parameters that will serve as a basis for us to, depending on the results, modify the following ones to find a balance that achieves the most optimal results.

### 5.3.1 Baseline Scenario

To gain insight into the behavior of our system, we initially established a set of baseline parameters. These parameters served as our starting point, enabling us to obtain preliminary results. Subsequently, during the evaluation process, we made incremental changes to a single parameter within each set, while keeping the remaining parameters at their baseline values. Table 5 illustrates the baseline simulation and the corresponding base values for each parameter.

| **Variable** | **Total** |
|---|---|
| Number of BNS | 64 |
| Number of anchors | 1 at heart |
| Granularity | 1/3 s |
| Event threshold detection | 1cm |

Table 5: Baseline scenario

To ensure meaningful and dependable results, we have carefully selected 64 nanonodes for deployment in our simulations. This number strikes a balance between avoiding insignificance and unreliability that a smaller sample size may produce, while also steering clear of an excessively high number. Our primary objective in this initial stage is to validate the proper functioning of our system and investigate its behavior. As for the number of anchors, we begin with the minimum requirement, as previously mentioned, which involves placing an anchor in the heart/thorax region. Additionally, we will initiate the simulations with a sampling rate of 3 samples per second and a 1cm event threshold for detection.

# 6 Experiments and Results

Each evaluation scenario within a set will present a figure comprising three types of graphs. Firstly, we include a graph that assesses accuracy in relation to simulation time, this graphs represents the accuracy (positive detection/total detection) versus the time. Secondly, we provide a graph displaying the point accuracy, which is the error in distance for each region. This graph allows us to observe the average error in distance for predictions, as well as identify any outlier values. To best represent this data, we have chosen to utilize a boxplot, as it offers a clear visualization of the desired results. Lastly, the third graph illustrates the most frequently predicted region, which can also be interpreted as a measure of accuracy. However, instead of calculating this for each simulation time, we have employed the mode to determine the most predicted region. We gathered all the predictions for each simulated region and calculated the mode to identify the region with the highest frequency of predictions. Subsequently, we compared this most predicted region against the actual region. A value of 1 indicates a match between the most predicted and actual regions, while a value of 0 indicates a mismatch.

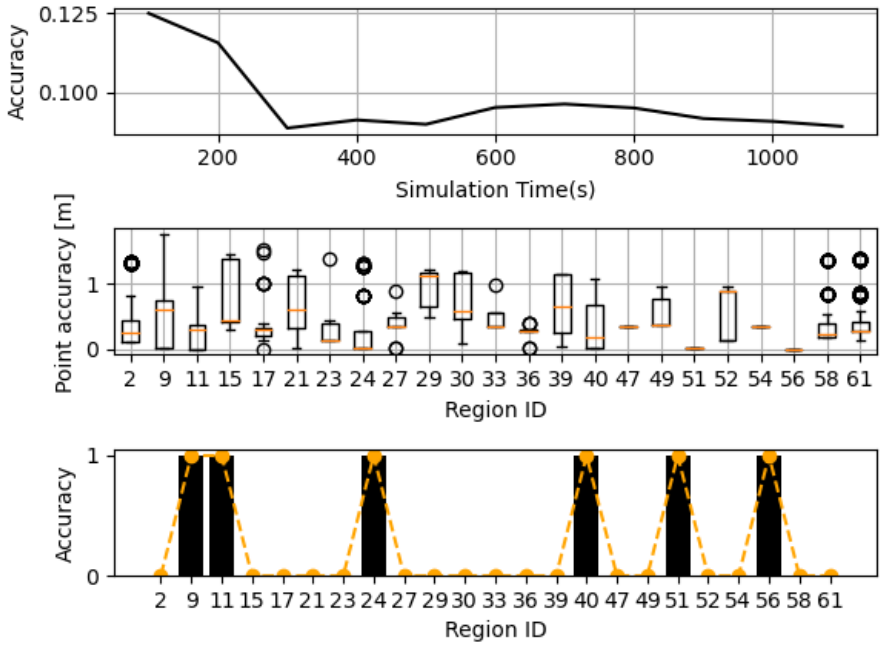
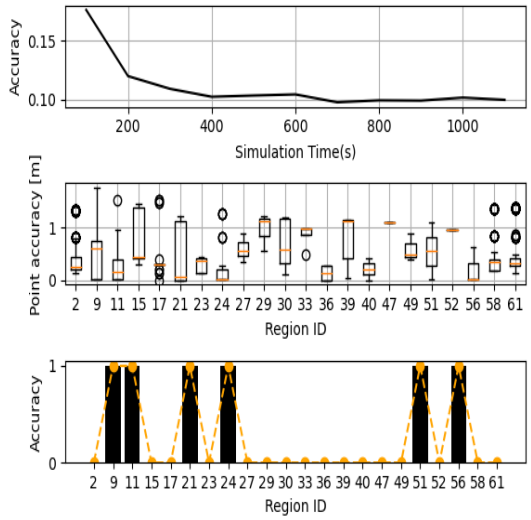



Figure 11: Baseline evaluation

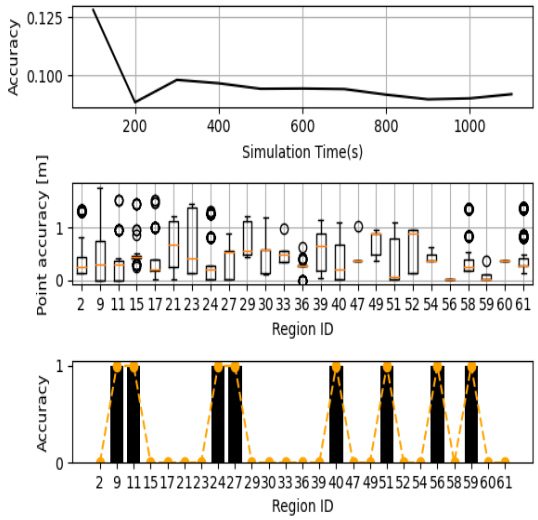
Firstly, we will present the base scenario that was previously discussed. The figure above showcases the three graphs mentioned earlier. In the first graph, we observe that the accuracy remains below 10% throughout the simulation period. Moving on to the second graph, we can observe the variation in distance errors across different regions. It is evident that only regions 51 and 56 have a mean error of 0, indicating accurate predictions. However, it is important to note that most regions exhibit outlier values, which is a

common occurrence in simulations of this nature due to communication interference and insufficient energy for reliable data transmission. Lastly, the third graph demonstrates that when we consider the mode of predicted regions for each actual region instead of analyzing individual samples, only 6 out of the 25 regions align with the original regions.

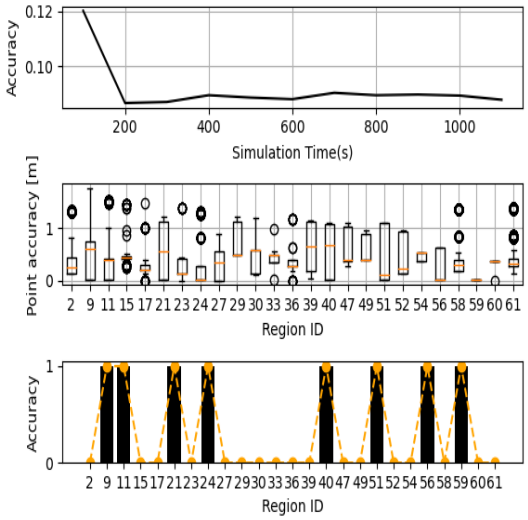
## 6.1 Scaling with the Number of Nanonodes




(a) 32 nanonodes evaluation




(b) 128 nanonodes evaluation




(c) 250 nanonodes evaluation

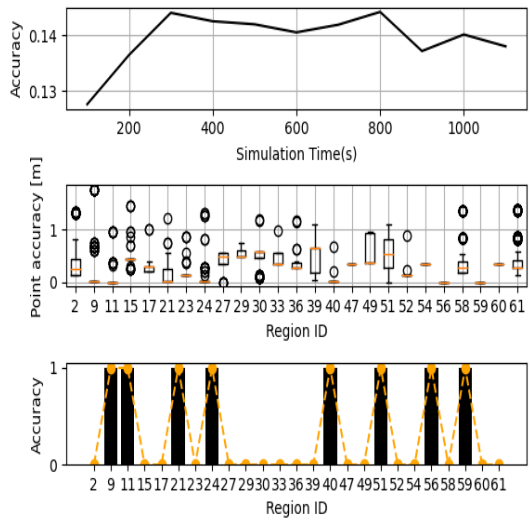
Figure 12: Localization performance as a function of the number of nanonodes

In this set of scenarios, we assessed the behavior of our system by manipulating the number of nanonodes within the body. It is evident that as we increase the number of nanonodes, there is a slight decrease in accuracy, as depicted in the first graph. Moving on to the second graph, we observe a slight decrease in error distance; however, it is accompanied by an increase in the number of outlier values. Finally, in the third graph, we can see a rise in the most predicted value per region. Consequently, we observe a greater number of regions where the mode (most frequently occurring value) aligns with the correct region, reaching a maximum of 8 regions.

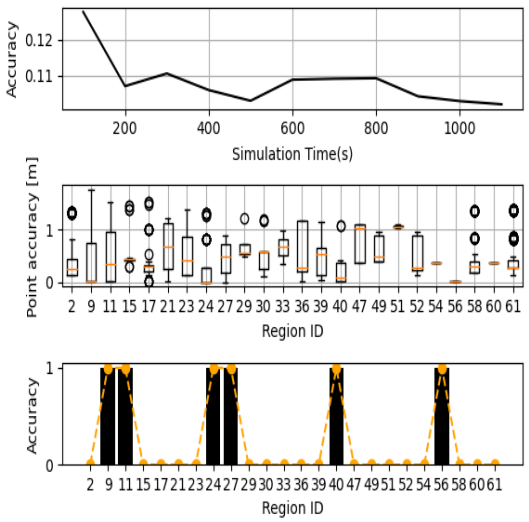
The results depicted in the first graph may appear somewhat counterintuitive at first, but it is logical that the accuracy experiences a slight decline when the number of nanonodes increases. This can be attributed to the initial implementation of our ML model, which tends to produce more incorrect predictions than correct ones. Consequently, while we may also generate more true positives, the number of negatives increases even more, resulting in a decrease in overall accuracy. However, upon examining the third graph, we observe the advantages and improved outcomes of calculating accuracy using the mode. This approach allows us to predict regions more accurately, leading to better results overall.

In conclusion, increasing the number of nanonodes leads to improvements in the number of predicted regions which allow to predict two more regions correctly in contrast with the baseline scenario and it also reduces the error in distance. However, it results in a decrease in accuracy in relation to the simulation time, as we have more samples miss predicted.

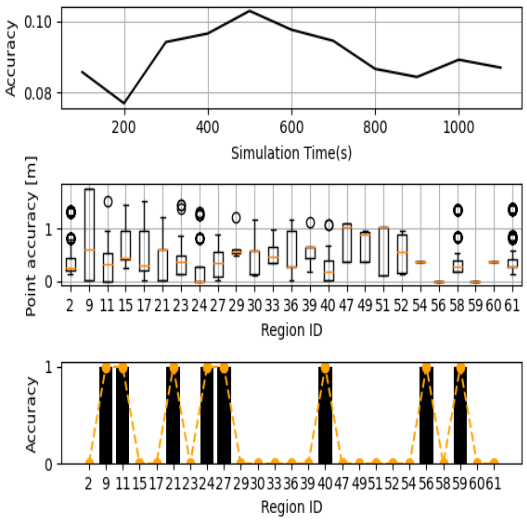
## 6.2 Scaling with Granularity



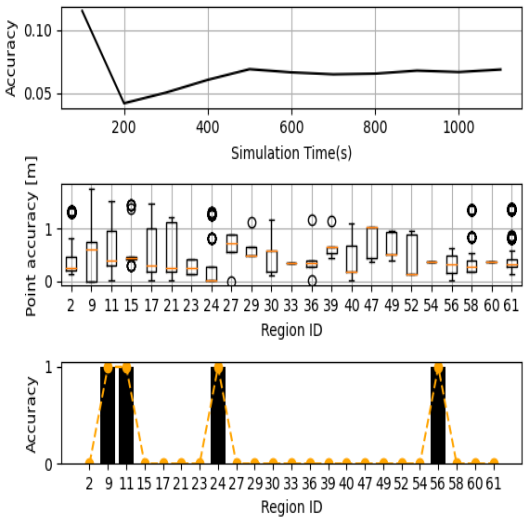

(a) 1 sample/s evaluation




(b) 2 samples/s evaluation




(c) 5 samples/s evaluation




(d) 10 samples/s evaluation

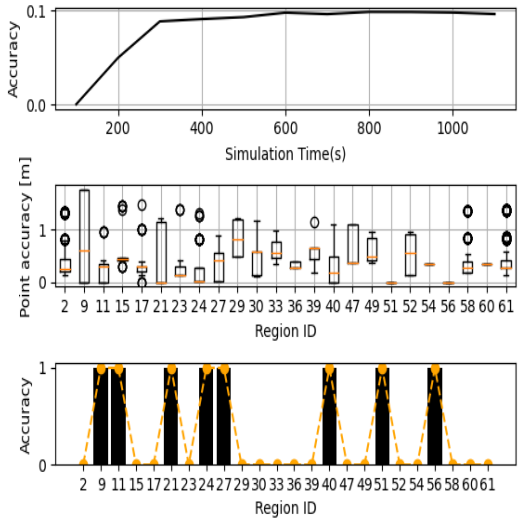
Figure 13: Localization performance as a function of temporal granularity

In this series of simulations, we have examined the behavior of our system by manipulating the frequency of upsampling. Notably, we have observed a decrease in accuracy as the

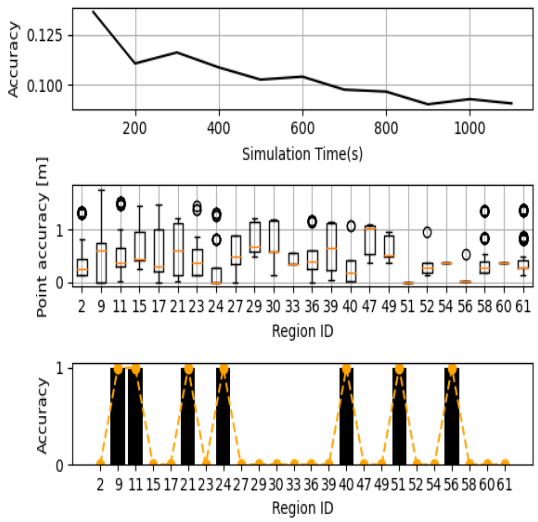
frequency increases. Furthermore, an interesting observation is that the most predicted region tends to decrease as the frequency is raised.

Initially, during the evaluations, there was an expectation that increasing the number of samples per second would result in a higher number of regions being predicted correctly. However, upon conducting the evaluations, it became evident that this assumption did not hold true. The reason behind this lies in the behavior of our system. While an increase in the number of samples does generate more positive predictions, it also introduces an elevated number of negative predictions due to noise and reduced precision in certain regions. As a result, both the precision in relation to circulation time and the number of correctly predicted regions decrease when the frequency is increased.

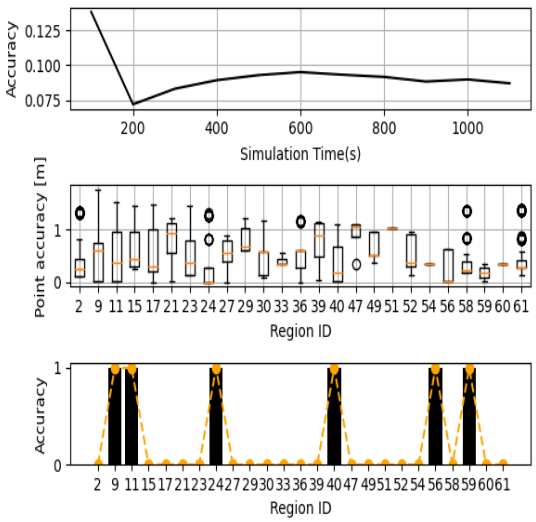
## 6.3 Scaling with the Event Threshold




(a) 0.5 cm event threshold




(b) 2 cm event threshold




(c) 3 cm event threshold

Figure 14: Localization performance as a function of the event threshold

In this set of scenarios, we examined the behavior of our system by manipulating the threshold of detection for each event within the body. The first graph illustrates the accuracy, which experiences a slight decrease as the threshold is increased. Notably, the best performance in terms of accuracy versus time is observed when the threshold is set at

0.5cm, maintaining a consistent 10%. As we further increase the threshold, the accuracy continues to decrease, gradually. Moving on to the second graph, we observe a decrease in the mean error in distance for certain regions as the threshold is adjusted. This indicates an improvement in prediction accuracy for those specific regions. Lastly, the third graph highlights the relationship between the threshold and the number of correctly predicted regions. Lowering the threshold results in a higher number of regions being predicted correctly. Conversely, as the threshold is increased, the number of correctly predicted regions decreases.

In theory, one might expect that increasing the threshold would lead to the detection of more events, thereby improving the overall accuracy. However, as observed in graph 3, the number of correctly predicted regions actually decreases as the threshold is raised. This unexpected outcome can be explained by the fact that, as we previously mentioned, the events in the evaluations were randomly located within the region rather than being centralized. Consequently, when the threshold is increased, an event may be detected even if the corresponding nanonode is not in the same region as the event itself. This discrepancy results in a higher number of erroneous predictions, contributing to the decline in accurate region predictions.

These results serve as a valuable reminder of the significance of the assumptions we make before conducting evaluations. As demonstrated in this set of simulations, our assumptions can have a substantial impact on the anticipated outcomes. Therefore, it is crucial to carefully consider and validate our assumptions to ensure the reliability and relevance of the expected results.

# 7 Conclusions and Future Work

As you have seen throughout this thesis, our work offers a model that serves as the foundation for standardized evaluation of future solutions. This allows all researchers to work hand-in-hand, with a correct and solid foundation to start from, and continue advancing in this area that has so many benefits to offer us.

Building upon our successful development of a highly accurate simulator for replicating the internal environment of the human body, we have achieved a significant milestone by establishing stable and reliable two-way communication between nanonodes and anchors. This achievement has paved the way for a standardized model that facilitates event-based communications within the human body, effectively transcending the inherent constraints imposed by both the human body and the nanonodes. This breakthrough brings us closer to unlocking a deeper understanding of communication dynamics within the human body and opens up exciting possibilities for advancements in healthcare technologies and remote monitoring.

In addition to the successful development of the simulator, we have also implemented a carefully designed standardized evaluation methodology for the flow guided localization solution under consideration. The purpose of this methodology is twofold: to validate the accurate functioning of the simulator and to gain valuable insights into the operation of our system. These crucial initial steps provide us with a solid foundation for future enhancements and advancements in our research and development endeavors.

The results indicate that the proposed workflow and the simulator can be utilized for capturing the performance of flow-guided localization approaches in a way that allows objective comparison with other approaches. Our results reveal relatively improvable accuracy of the different evaluated scenarios.This is due to unreliable THz communication between in-body nanonodes and on-body anchors and intermittent operation of the nanonodes due to energy-harvesting. It has also been shown in the evaluations how some regions are never detected, this is because the circulations times in some regions, most of them on the upper body (right and left heart, thorax, intestine, hips) their circulations times are too close between them, and some minimum delay in the transmission, makes miss the prediction, also the lack of data make also hard the training of some regions. There is other problems regarding the predictions in some regions for example when trying to distinguish between a region of the body and its pair region (for example between left and right hand), for our actual system it is very hard to distinguish between them if their circulation times are similar or equal.

We can observe that the accuracy obtained is relatively low, which is our most important limitation that we must focus on. We also have limitations in terms of scalability, requiring further code and system structure modifications. Another limitation, although it is more a problem, would be the disparity of data obtained in different simulations, which hinders the training of our ML model and directly affects the reliability and accuracy of our results.

As we mentioned in the previous paragraph, our priority for the future is to improve the levels of accuracy we have obtained in this thesis. Clearly, much higher values are needed

to consider them reliable, so this must be the main focus in the future. Modifications will need to be made to both the programming and structure of the designed system to achieve greater scalability in the future. Furthermore we will need to continue improving the proposed ML model. As we have mentioned, it will need better tools to facilitate/improve its training, with that we should be able to predict some regions that now are not predicted well due to their similarities in their simulation times, and with that obtain better results. It is also important to make some improvements for solve the prediction problems between the regions that has two pairs equally located (for example right and left hand). Finally, there is also a need to evaluate this system with more than one anchor as explained in chapter 5 due to for some bugs regarding the ML algorithm the evaluations could not be done, although our simulator is prepared to work with more than one anchor.

Part of the work presented in this thesis has been submitted as a paper to the IEEE Communication Standards magazine. Specifically, the section on the development of the simulator and evaluation of it has been selected for publication. This paper [15] provides a detailed description of the fundamentals, design, and testing of the proposed flow guided localization solution, as well as the performance evaluation. The paper also includes a discussion of the potential applications of these devices in the detection of health problems and drug delivery. The possible acceptance of this paper for publication in the IEEE magazine is a significant accomplishment and highlights the contribution of this work to the field of nanotechnology and wireless communication.

# 8 Work Plan

We tried to design a realistic work plan, with a clear split of the different task that we should develop from the beginning until the end. For every task we estimated a time schedule, based on what we though every task would spend. While developing the tasks we encounter some issues, specifically in the development of the simulator, we ended up having some unforeseen problems, in the development of the communication between the anchor and the nanonode. Understanding where this issues were and how to solve them makes us spend more time than what we thought we would spend. Another problem we had was developing the in-body propagation loss, recreating the in-body environment with all the constraints and attenuation the in-body has, to adapt that to our simulation. Also, the data processing was also something more complicated that what I initially thought due to the amount of data generated that needs processing.

These problems mentioned above have caused us to delay the completion of our work and we were unable to deliver the complete thesis on time, so we had to use a little more time to finish polishing and fixing our work. The tables below are detailing the changes made in our work plan in this final steps of our thesis.

<table>
<tr><td colspan="2">Project: Coarse-grained Localization of In-body Energy-harvesting Nanonodes</td><td colspan="2">WP ref (WP1)</td></tr>
<tr><td colspan="2">Principal task: Research and study</td><td colspan="2">Table 1 of 5</td></tr>
<tr><td colspan="2" rowspan="2">Description:<br>Read different published papers about THz nanonetworks and in-body localization.</td><td colspan="2">Planned start date: 01/09/2022<br>Planned end date: 30/09/2022</td></tr>
<tr><td colspan="2">Start event: 01/09/2022<br>End event: 30/09/2022</td></tr>
<tr><td colspan="2">Internal task T1.1. Read and research papers about THz nanonetworks.<br><br>Internal task T1.2. Read and research papers about in-body localization.</td><td>Deliverables:<br>N/A</td><td>Dates:<br>N/A</td></tr>
</table>

Table 6: Work package 1

<table>
<tr><td colspan="2">Project: Coarse-grained Localization of In-body Energy-harvesting Nanonodes</td><td colspan="2">WP ref (WP2)</td></tr>
<tr><td colspan="2">Principal task: Design space specification for flow-guided THz localization</td><td colspan="2">Table 2 of 5</td></tr>
<tr><td colspan="2" rowspan="2">Description:<br>Design the parameters needed for flow-guided THz localization.</td><td colspan="2">Planned start date: 07/10/2022<br>Planned end date: 29/10/2022</td></tr>
<tr><td colspan="2">Start event: 3/10/2022<br>End event: 15/10/2022</td></tr>
<tr><td colspan="2">Internal task T2.1. Design the parameters for flow-guided Thz localization.<br><br>Internal task T1.2.</td><td>Deliverables:<br>N/A</td><td>Dates:<br>N/A</td></tr>
</table>

Table 7: Work package 2

| Project: Coarse-grained Localization of In-body Energy-harvesting Nanonodes | WP ref (WP3) | |
|---|---|---|
| Principal task: Implementation of simulation setup and designed localization solution | Table 3 of 5 | |
| Description:<br>Developing the simulator with Blood voyager-s, NS-3 and teraSim.<br>Make all the intermediate scripts to process the logs generated. | Planned start date: 29/10/20222<br>Planned end date: 24/11/2022<br>Start event: 17/11/2022<br>End event: 17/03/2023 | |
| Internal task T3.1. Setup of the differents softwares needed for the implementation<br><br>Internal task T3.2. Develop our simulation with our designed space specifications.<br><br>Internal task T3.3. Generate the logs and generate scripts to process them | Deliverables:<br>1.The simulation developed | Dates:<br>N/A |

Table 8: Work package 3

| Project: Coarse-grained Localization of In-body Energy-harvesting Nanonodes | WP ref (WP4) | |
|---|---|---|
| Principal task: Design space exploration and analysis | Table 4 of 5 | |
| Description:<br>Design and implement the flow guided localization solution. | Planned start date: 05/12/2022<br>Planned end date: 24/12/2022<br>Start event: 18/03/2022<br>End event: 01/05/2023 | |
| Internal task T4.1.Implement the flow guided localization solution.<br><br>Internal task T4.2. Generate the evaluation scenarios. | Deliverables:<br>Results and conclusions | Dates:<br>N/A |

Table 9: Work package 4

| Project: Coarse-grained Localization of In-body Energy-harvesting Nanonodes | WP ref (WP5) | |
|---|---|---|
| Principal task: Implementation of simulation setup and designed localization solution | Table 5 of 5 | |
| Description: Write and defend the thesis . | Planned start date: 12/12/2022<br>Planned end date: 10/01/2023<br>Start event:01/05/2023<br>End event:13/05/2023 | |
| Internal task T1.1. Write the thesis<br><br>Internal task T1.2. Design the final presentation of the thesis | Deliverables:<br>N/A | Dates:<br>N/A |

Table 10: Work package 5

## 8.1 Gantt Diagram

According to the explained before at the work plan, we had some issues that delayed the development of our thesis. Obviously, we also had to make changes in the Gantt Diagram to show that issues. Now we are representing the final Gantt Diagram of our thesis.

| *Updated Time Plan (Gantt diagram)* | August | | September | | October | | November | | December | | | January | | February | | MARCH | | APRIL | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 25-1 | 1-15 | 15-29 | 29-13 | 13-27 | 27-10 | 10-24 | 24-08 | 8-22 | 22-05 | 5-19 | 20-2 | 2-16 | 17-3 | 3-16 | 17-30 | 31-13 | 14-28 |
| WP1: Research and study | | | | | | | | | | | | | | | | | | | |
| Internal Task 1: Read and research papers about nanonetworks | | | | | | | | | | | | | | | | | | | |
| Internal Task 2: Read and research papers about in-body localization and communication | | | | | | | | | | | | | | | | | | | |
| WP2: Design space specification for flow-guided THz localization | | | | | | | | | | | | | | | | | | | |
| Internal Task 1: Design the parameters for flow-guided Thz localization. | | | | | | | | | | | | | | | | | | | |
| WP3: Implementation of simulation setup and designed localization solution | | | | | | | | | | | | | | | | | | | |
| Internal Task 1: Setup of the differents softwares needed for the implementation | | | | | | | | | | | | | | | | | | | |
| Internal Task 2: Develop our simulation with our designed space specifications | | | | | | | | | | | | | | | | | | | |
| Internal Task 3: Design the different scenarios | | | | | | | | | | | | | | | | | | | |
| Internal Task 4: Generate the raw data for every scenario designed and process that data | | | | | | | | | | | | | | | | | | | |
| WP4: Evaluation and analysis | | | | | | | | | | | | | | | | | | | |
| Internal Task 1: Extract all the relevant results and metrics. | | | | | | | | | | | | | | | | | | | |
| Internal Task 2: Analyze them and extract conclusions of the system performance. | | | | | | | | | | | | | | | | | | | |
| WP5: Project Documentation | | | | | | | | | | | | | | | | | | | |
| Internal Task 1: Write the project final report. | | | | | | | | | | | | | | | | | | | |
| Internal Task 2: Create a template for the oral presentation. | | | | | | | | | | | | | | | | | | | |

Figure 15: Gantt diagram

# References

[1] J. M. Jornet and I. F. Akyildiz. Joint energy harvesting and communication analysis for perpetual wireless nanosensor networks in the terahertz band. *IEEE Transactions on Nanotechnology*, 11(3):570–580, 2012.

[2] Q. H. Abbasi K. Yang N. Chopra et al. Nano-communication for biomedical applications: A review on the state-of-the-art from physical layers to novel networking concepts. *IEEE Access*, 4:3920–3935, 2013.

[3] F. Lemic et al. Survey on terahertz nanocommunication and networking: A top-down perspective. *IEEE Journal on Selected Areas in Communications*, 39(6):1506–1543, 2021.

[4] S. Abadal I. Llatser A. Mestres et al. Time-domain analysis of graphene-based miniaturized antennas for ultra-short-range impulse radio communications. *IEEE Transactions on communications*, 63(4):1470–1482, 2015.

[5] F. Dressler and S. Fischer. Connecting in-body nano communication with body area networks: Challenges and opportunities of the internet of nano things. *Nano Communication Networks*, 6(2):29–38, 2015.

[6] F. Lemic S. Abadal A. Stevanovic E. Alarcon and J. Famaey. Toward location-aware in-body terahertz nanonetworks with energy harvesting. *in 9th ACM International Conference on Nanoscale Computing and Communication*, pages 01–06, 2022.

[7] J. Simonjan B. D. Unluturk and I. F. Akyildiz. In-body bionanosensor localization for anomaly detection via inertial positioning and thz backscattering communication. *IEEE Transactions on NanoBioscience*, 21(2):216–225, 2021.

[8] R. Geyer M. Stelzner F. Buther and S. Ebers. Bloodvoyagers: Simulation of the work environment of medical nanobots. *in 5th ACM International Conference on Nanoscale Computing and Communication*, pages 1–6, 2018.

[9] Z. Hossain Q. Xia and J. M. Jornet. Terasim: An ns-3 extension to simulate terahertz-band communication networks. *Nano Communication Networks*, 17(2):36–44, 2018.

[10] T. Van Haute et al. Platform for benchmarking of rf-based indoor localization solutions. *IEEE Communications Magazine*, 53(9):126–133, 2015.

[11] N. Moayeri et al. Perfloc (part 1): An extensive data repository for development of smartphone indoor localization apps. *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), IEEE*, page 1–7, 2016.

[12] D. Lymberopoulos et al. A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned. *in IEEE/ACM International conference on information processing in sensor networks*, pages 178–189, 2015.

[13] J. T. Gomez A. Kuestner J. Simonjan B. D. Unluturk and F.Dressler. Nanosensor location estimation in the human circulatory system using machine learning. *IEEE Transactions on Nanotechnology*, 21(2):663–673, 2022.

[14] D. Vasisht G. Zhang O. Abari H.-M. Lu J. Flanz and D. Katabi. In-body backscatter communication and localization. *in ACM Special Interest Group on Data Communication*, pages 132–146, 2019.

[15] Brosa Lopez A Lemic F Struye J Torres Gómez J Municio E Delgado Pinillos C Dressler F Alarcón E Famaey J Abadal S Costa Perez X. Toward standardized performance evaluation of flow-guided nanoscale localization. *IEEE Communications Standards Magazine*, 2023.

If you want more details about the implementation of the simulator, the repository with the system code is attached:

https://bitbucket.org/filip_lemic/flow-guided-localization-in-ns3/src/master/

It will also be attached a zip of all the simulator in the annex.