Nanosensor Location Estimation in the Human Circulatory System using Machine Learning

Jorge Torres Gómez, Senior Member, IEEE, Anke Kuestner, Jennifer Simonjan, Member, IEEE, Bige Deniz Unluturk, Member, IEEE, and Falko Dressler, Fellow, IEEE

Abstract—The human body can be considered a complex natural network due to the variety of interconnections between the different body regions. One example is the network of blood vessels, where artificial communication channels can be rendered using nanosensors that travel in the bloodstream as collectors and carriers of information. Further advancing this vision, in this work we investigate the detection and localization capabilities of flowing nanosensors in the blood flow to report abnormalities in the human body. Specifically, we target the detection of quorum sensing molecules and provide a methodology to evaluate its performance. The methodology consists of modeling the traveling path of nanosensors along the vessels through a Markov chain, and the use of machine learning (ML) models to compute their transition probabilities. We illustrate the resulting distribution of nanosensors in the body, which evidences a close match to expected results. We also evaluate their detection and localization capabilities in different body regions revealing their effectiveness to determine the presence of abnormalities in the human vessels.

Index Terms—Nanosensors, Nano Communication, Human Circulatory System, Precision Medicine, Machine Learning, Markov Model

I. INTRODUCTION

FOLLOWING the concept “the body is the network” [1], the nature of the human circulatory system inherently provides such a “connectivity” as they interconnect all parts of the human body. Nanosensors injected into the blood flow are envisioned to collect and communicate data from different body regions supporting the detection and localization of abnormalities [2]. Nanosensors [3], [4] and their capabilities to communicate inside the body [5]–[7] are considered critical enabling technologies to conceive such a monitoring system in vessel’s networks. This system will pave the way for not only health monitoring applications but also targeted drug delivery applications.

Novel nanomaterials such as graphene and its derivatives have enabled the fabrication of tiny electronic devices including sensors, processors, memories and transceivers [8]. Smart nanosensors [4] can be enabled to record small concentration of biological and chemical molecules, or local changes in physical variables, such as temperature, pressure, or vibrations [9] along the vessels. Envisioning the smart nanosensor concept, the sensor readings can then be processed, stored or communicated to an external gateway for the later processing.

As one potential scenario, being inside the body and having a significantly higher sensitivity, the nanosensors can detect much lower concentrations enabling the detection of diseases at a much earlier development stage [10], [11]. Nanosensors technology, although today in the early development stage, is further devised for in vivo testbeds for its later use in healthcare applications [12]. In general, it takes a long time for biomarker molecules to reach a concentration high enough to be detected by standard sensing technology, often leading to late diagnosis of certain diseases such as cancer.

Further advancing this vision, in this work we study the abnormality detection capabilities of nanosensors, which are flowing in the HCS as depicted in Fig. 1 (a). As a potential application scenario, we consider the case study where nanosensors are used to detect infection biomarkers released by bacteria, as depicted in (b). Then, they report their findings to an external Gateway (fusion node) when traveling through the left heart, as depicted in (c).

For this scenario, we introduce a methodology to evaluate the detection capabilities of the flowing nanosensors. Through this methodology, we provide the resulting probability of detection and the location of the reported abnormalities in the human body, i.e., the given tissue. To compute the detection probability, we model the nanosensors traveling in the human vessels...
according to a Markov chain model [13], which in turn let us evaluate the concentration of nanosensors per vessel segment, and the resulting detection capabilities.

To estimate the location of abnormalities, we use a machine learning (ML) model at the Gateway (cf. Fig. 1 c)). This way, we can avoid the need for nanosensors to keep track of their own positions at all times which is infeasible for several reasons: Such an inertial localization approach would require sensors to be equipped with highly precise accelerometers and gyroscopes. Even if those were available, it would most likely be impossible to track positions accurately as nanosensors flow through a highly complex network of many branches and rotate around their own axes. Additionally, tracking their positions in a timely manner would require them to gather and store huge amounts of sensor data which can barely be achieved on such resource-constrained devices.

Furthermore, the use of ML models provides two additional benefits: On the one hand, it avoids the complexities of modeling the human vessels [14] and their fluid dynamics (laminar flow and turbulence) [15]. On the other hand, it is inherently adaptable to the variety of physiological parameters in the human body. The ML models we propose will simply use the traveling times and concentration levels of nanosensors to predict the origin of reported detections.

In this work, building upon our earlier work in [16], we include a new dimension to train the ML. We no longer rely on the travelling times to predict the distribution of nanosensors only, but simultaneously exploit the concentration level of nanosensors per vessel segment as provided by the BloodVoyagerS (BVS) simulation framework [17]. This new dimension of data further improve the localization capabilities in the system. In this way, our main contributions can be summarized as follows:

- We introduce a methodology to effectively evaluate the detection capabilities of nanosensors flowing in the HCS. The methodology consists of a Markov chain process, where its transition probabilities are researched using ML methods.
- We introduce the use of low-complex ML models to predict the location of abnormalities reported by nanosensors. In this work specifically, we also include one additional dimension aiming to further improve the detection performance.
- Following the recommendation from Gebru et al. [18], we provide data and documented it for the training of the machine learning model,¹ and results illustrating the concentration level of nanosensors along the vessels.

The rest of the paper is structured as follows. In Section II, we provide an overview of the different techniques for in-body localization of nanosensors and detection of abnormalities. Section III introduces the local detection mechanism of quorum sensing molecules and the methodology to report detected diseases. The ML model is introduced in Section V. The evaluation and results are presented in Section VI, and we conclude the paper and outline future work in Section VII.

II. RELATED WORK

Being able to localize the nanosensors’ measurements is crucial to localize detected abnormalities in the given body region. Localization in a challenging environment such as the HCS, where nanosensors are highly mobile, constrained and with limited communication capabilities, is fairly difficult. Therefore, conventional localization or tracking systems fail and novel solutions are required [19], [20]. To the best of our knowledge, there are only very few approaches attempting to study the localization of in-body nanosensors or detected abnormalities. Additionally, most of the attempts focus on localizing the sensors themselves rather than the detected abnormalities, which is very difficult in reality to achieve due to the high mobility and limited communication capabilities of nanosensors.

In the literature, there are several methods which assume an abstract network topology or known sensor locations. One example is the work of Mosayebi et al. [21], who propose to use mobile nanosensors for detection of cancer at early stages where cancer cells are located in particular regions of a blood vessel. Their focus is mostly on the biomarker detection capabilities as they assume a known network topology of 16 edges and a total of 96 cm for all vessels. Varshney et al. [22] address the abnormality detection problem in a blood vessel using multiple cooperative nanosensors and a common receiver called the fusion center (both mobile). Sensors perform abnormality detection which is associated with probability of detection and probability of false alarm, and their local decisions are reported to the fusion center. The authors assume an abstracted network topology and perfect time synchronization between all nodes. Meanwhile, a theoretical framework for cooperative abnormality detection and localization systems, exploiting a molecular communication setup, was proposed by Khaloupour et al. [23]. Their system constitutes of mobile sensors in a fluid medium in which the sensors search for abnormalities. Additionally, fusion centers are deployed at various locations to collect and fuse all sensor measurements. Each fusion center covers a specific location which can have at most one abnormality leading to a very coarse grained localization resolution.

Another group of related works focus on the localization of the nanosensors themselves, which is computationally complex and realistically difficult to achieve in the HCS environment. Additionally, all of the approaches are tested with very general simulation setups which can barely evaluate the in-body scenario. One example is the approach of Lemic et al. [24] which introduced an in-vivo localization method for nanosensors that exploit wireless nanocommunication techniques based on Wake-up Radio (WuR) and Software-Defined Metamaterials. To evaluate this approach, they simulated a circular area of 30 cm with randomly deployed nodes. Another nanosensor localization algorithm was proposed by Zhou et al. [25]. It utilizes pulse-based distance accumulation to determine distances between anchors and clustered nodes. Flooding is used to cluster the

¹The data is generated after post-processing the coordinates from traveling nanosensors using the BVS simulator [17]. The code and the data is publicly accessible at https://github.com/jorge-torresgomez/BVS_data
networked nodes first into corner, border and center nodes, followed by a clustering algorithm to determine cluster heads. Distances between nodes are then estimated by cluster heads based on the hop-count of messages. Similar to the evaluation scenario above, the authors considered randomly deployed nodes on a 2D square area which can barely evaluate the in-body environment.

One of our previous works tackles abnormality detection with a similar idea [26]. The authors focus on the localization of abnormalities, which are detected by nanosensors flowing through the HCS. To do so, the system requires multiple anchors attached to the human body and the nanosensors need to be equipped with an inertial measurement unit as they exploit inertial positioning for self-localization purposes. For realistic evaluations of the bloodflow, the authors also exploited BVS [17].

Compared to existing research, the approach presented here is a completely novel concept towards evaluating the detection performance and inferring the location of nanosensors in the HCS based on ML models [16]. Besides the concept for detection and localization, realistic simulation scenarios enable to evaluate the approach for an actual in-body scenario. This approach provides a widely applicable solution to in-body nanosensor simulations without depending on the specifics of the physiological parameters (e.g., lengths of vessels, blood viscosity and pressure [27]), and the complex blood flow regimes in the vessels (e.g., turbulences or laminar flows [28]).

III. MECHANISM TO REPORT THE DETECTION OF INFECTIOUS DISEASES

To elaborate on the detection mechanism, the mobile nanosensors will detect infectious diseases by eavesdropping to the natural communication among bacteria, which is called quorum sensing (QS). Among the bacterial cells, QS molecules are exchanged for communication and the amount of QS molecules is correlated with the bacterial population density. During the infection by bacteria, their population density increases and hence more QS molecules are produced. The mobile nanosensors will sense the concentration of QS molecules. If it is above a certain threshold, it indicates that the bacterial population is big enough to produce infections [29]. The human body inherently contains many bacteria in large numbers which are harmless and even helpful for some body functions such as digestion and protection against pathogens. However, when pathogenic bacteria infect the body and start growing out of control, serious health conditions may occur. To that end, detecting infections in the early stages allows timely response from clinicians and early administration of antibiotics, which leads to reduced symptoms and improved survival rate. However, conventional methods of infection detection involve culturing of bacteria to increase their quantity to detectable levels, which typically requires 48–72 hours. While alternative methods such as enzyme linked immunosorbent assay (ELISA) and polymerase chain reaction (PCR) provide higher sensitivity and specificity within a shorter assay time, they require complex instrumentation and skilled operators limiting their use to clinical laboratories. Hence, these methods are not suitable for continuous in vivo monitoring.

QS is a crucial mechanism in the onset of infection. It allows individual bacteria to collaborate towards forming a biofilm where cells lie in an extracellular matrix further protecting them from body’s natural defenses. During this process, QS molecules are produced in large quantities and eventually they diffuse from the infection site into the bloodstream. The level of QS molecules reaches detectable levels in all bodily fluids such as blood, saliva, and sputum in correlation with the progress of infection [30], [31]. Nanosensors injected into the bloodstream are considered to be capable of sensing these QS molecules to continuously monitor infections. Examples of such injectable sensors used for various disease diagnosis and therapeutics are electronics based sensors such as neural dust which are grain sized wireless brain recording devices [32] and syringe-injectable electronics [33]. Cell-based examples of injectables are programmable probiotic bacteria for cancer detection [34] and microsensors circulating in the bloodstream mimicking red blood cells for sodium [35]. The following subsections discuss the local detection mechanism and the methodology to report the detection of infectious diseases.

A. Local detection mechanism

To measure the QS molecules concentration, mobile nanosensors can utilize either electrochemical sensors such as the ones measuring QS molecules of Pseudomonas aeruginosa pyocyanin [36] and N-Acyl homoserine lactones (AHL) [29] through electrodes and converting them to current signals.

The distribution of QS molecules near an infection site is simulated using COMSOL. Using our previous study [37], we simulate the distribution of QS molecules in the tissue following diffusion in porous medium where the interstitial space is considered as pores through which QS molecules are diffusing. We consider a representative tissue volume where the infection site is on top and the capillaries are at the bottom of this unit volume. The concentration of QS molecules is shown in Fig. 2 for the time instant bacteria population reaches critical density for switching to pathogenic behavior.

Following the findings in [37], it is possible to determine the region where the concentration of QS molecules is above the threshold that the nanosensors can detect, which we call sensing region. If the detection limit of nanosensors is considered to be $1 \times 10^{-5}$ million [38], an ellipse can be used to approximate the sensing region which is elongated in the blood flow direction with axes lengths of 0.1–0.5 cm and 1.75–3.5 cm.

Given the region of interest of volume $V_k$ (e.g., capillaries in the legs, center body, and arms), the conditional probability for a nanosensor to visit the sensing region $V_{s,k} \subset V_k$ can be expressed as

$$P_{s|k} = \text{Prob}(l_{n}^s \in V_{s,k}|l_{n} \in V_k) = \frac{V_{s,k}}{V_k},$$

where $l_{n}^s$ is the $(x, y)$-location of nanosensors, $V_{s,k}$ is the volume of the sensing region, determined through the COMSOL simulations for the organ $k$, and $V_k$ is the total volume for the given organ.

Most bacteria reproduce in a timeframe of 20 min, which is set as our observation time. This observation window is
set such that the concentration of QS molecules will not undergo significant changes and the nanosensors will complete several laps through the cardiovascular system, hence collecting multiple samples. Even if one of these samples are detecting infection, the nanosensor will report that it detected an infection somewhere on its route during the observation window.

Then, the probability of successful detection, denoted as $P_{d,k}$, is determined as the probability of at least $N_{th}$ nanosensors reporting sensed QS molecules as

$$P_{d,k} = \sum_{i=N_{th}}^{N_s} \binom{N_s}{i} P_s(i, P_{s,k} P_{c,k})$$  \hspace{1cm} (2)

$$= \sum_{i=N_{th}}^{N_s} \left(\frac{N_s}{i}\right) (P_{s,k} P_{c,k})^i (1 - P_{s,k} P_{c,k})^{N_s-i},$$

where $\binom{N_s}{i}$ denotes the binomial distribution, $N_{th}$ is the arbitrary threshold for the number of nanosensors needed to decide towards detection of infection, $P_{s,k}$ is given in Eq. (1), and $P_{c,k}$ is the probability of a nanosensor being located in the capillaries of organ $k$. In this equation, summing from $N_{th}$ to $N_s$ considers that at least $N_{th}$ sensors are visiting the given sensing region, which accounts for the probability of $i$ nanosensors visiting the sensing region out of a total of $N_s$. The two major factors used for evaluating this expression are given by $P_{s,k}$ and $P_{c,k}$, with the latter still to be determined.

**B. Methodology to report the detection of infectious diseases**

Fig. 3 provides a methodology to compute the binomial distribution according to Eq. (2). This methodology, to be implemented at the Gateway (cf. Fig. 1), aims to provide alerts with a given reliability level supporting clinical reports. The left branch concerns the evaluation of the visiting probability $P_{s,k}$, while the right branch the probability to find a nanosensor at the given capillary $k$. According to the left branch, we compute the $P_{s,k}$ term directly evaluating the expression in Eq. (1) and using the COMSOL simulation results to evaluate its numerator.

To estimate $P_{c,k}$, we exploit Markov chains and ML models as presented in the right branch of this methodology.

**IV. Modeling Schemes**

The dynamics of the nanosensors’ movement and their stationary distribution in the HCS are modeled by two primary schemes: the movement dynamics are provided by the BVS simulation framework [17] and the stationary distribution of nanosensors is derived via modeling the process as a Markov chain. The next two subsections provide further details on these two schemes.

A. Modeling the flow of nanosensors in the circulatory system

As the HCS comprises in total approx. 4900 cm$^3$ of blood volume and 120 000 km of blood vessels [39], modeling the whole system is very complex. However, to achieve a realistic movement model of the nanosensors traveling through the HCS, we use BVS [17], a nanonetwork simulation framework capable of simulating the blood flow of all major vessels in the HCS. BVS models a simplified version of the HCS to simulate the traveling behavior of particles in the human blood stream. All vessels and organs considered by the simulator are uniquely numbered and are depicted in Fig. 4. It comprises a model including 94 vessels and their respective blood flow rates (20 cm/s in the aorta, 10 cm/s in the arteries, 2–4 cm/s
in the veins), adding up to a total simulated vessel length of 12.717 m (vessel length measurements are based on a person with 1.72 m height and 69 kg weight).

BVS uses its own coordinate system with the origin in the heart (further details in [17]). Using BVS, we can thus collect the coordinates of all the travelling nanosensors over time. Based on these coordinates, we simulate the data delivery process to the Gateway when they are travelling through the right heart. As explained in Section V, from these coordinates we will evaluate the travelling time and the concentration level of nanosensors which are to be delivered to the ML module.

**B. Modeling the distribution of nanosensors as a Markov model**

Nanosensors inserted in the HCS will randomly travel through different trajectories according to the selected path at bifurcations in the arteries. For instance, considering the Arcus Aorta (A1), the blood flow will randomly bring the nanosensor to the head through the Ascendens (A4) or Carotid (A5) aortas, or to the center body through the Thoratica Aorta (A2). Assuming these random transitions at bifurcations are independent of the previously visited vessel, the process can be modeled according to a Markov chain [40].

To conceive the Markov chain, we establish a one-to-one correspondence between the different vessel segments provided by BVS and stages in the Markov chain, as depicted in Fig. 5. Through this model, a total of 30 nodes are defined for the arteries (A1 to A30), 25 nodes for the capillaries (C1 to C25), and 33 nodes for the veins (V1 to V33). In this way, the stationary probability to locate a nanosensor on a given vessel segment according to this model can be computed as

\[ P_{s,k} = \nu_k, \]

where \( \nu_k \) are the components of the stationary probability vector \( \nu \), obtained after solving the equation \( \nu = \nu P \), being \( P = \{p_{i,j}\} \) the transition matrix, and \( p_{i,j} \) the transition probabilities.

Following this procedure, the probability \( P_{s,k} \) depends on the transition matrix, and ultimately on the transition probabilities at the bifurcations \( p_{i,j} \). Intuitively, the transition probabilities at bifurcations can be directly evaluated according to the flow in vessels. For instance, the probability to jump from the Arcus Aorta (A1) to the Thoracic Aorta (A2) can be obtained as the ratio of the corresponding flows as

\[ P_{A1,A2} = \frac{I_{A2}}{I_{A1} + I_{A2}}, \]

where \( I_{A1} \) and \( I_{A2} \) denote the flow through the corresponding vessel segments Arcus and Thoratica aortas, respectively.\(^3\)

Furthermore, the relation in Eq. (4) can be estimated using the flow of nanosensors as they follow the blood flow [16]. The numerator can be evaluated when identifying the total of nanosensors travelling through the head from the total of nanosensors travelling through the Arcus Aorta, i.e., those nanosensors travelling through the loops L1 and L2 in Fig. 5. The denominator can be evaluated after identifying all the nanosensors travelling through the loops intercepting A1, i.e., L1, L2, L3 in Fig. 5 and the other loops closing at the left heart

\(^3\)the proof of this relation rely on the equivalent representation of the Markov model as closed circuits, were the probabilities are computed through the flows on those circuits, further details can be followed in our previous work in [13].
Human Circulatory System

\[ P_{A1,A2} = \frac{\sum_{i \in L_{A2}} N_i}{\sum_{i \in L_{A1}} N_i} \]  

where \( N_i \) is the total of nanosensors flowing on a given loop \( L_p \), and \( L_{A1} \) and \( L_{A2} \) represent the set of loops which pass through the vessel segments A1 and A2, respectively. In the next Section we will derive a ML-based mechanism to identify the total of nanosensors per loop, then to evaluate the transition probabilities of the Markov chain model.

V. LEARNING METHODOLOGY

In the system, the ML module runs at the external gateway using data provided by the flowing nanosensors (cf. Fig. 1). This module supports the computation of the transition probabilities from the Markov chain and generates alerts whenever QS molecules are detected. For that purpose, it consists of two parts: An unsupervised ML model to estimate the total of nanosensors per closed-loop (\( N_i \) in Eq. 5, cf. Fig. 3), and a supervised model to provide promptly alerts (detected QS molecules and their location). These two models are applied according to the four steps flow depicted in Fig. 6. Data to train the ML models are obtained from the BVS simulator (step 1), where nanosensors report their collected variables to the gateway. Then, the unsupervised ML model is used to cluster these data (step 2) supporting the functioning of the Markov model (step 3) on the one hand, and training the supervised ML model (step 4) on the other.

In our previous work [16], we explored the machine-learning detection performance exploiting only the travelling time reported by nanosensors. Intuitively, the nanosensors with longer traveling times are those located on the largest circuit paths (e.g., the legs), while the ones with shorter traveling times are coming from paths close to the heart (e.g., thorax). The reported time is then used to predict which circuit has been travelled by the respective nanosensor. Additionally, anchor nodes were used to distinguish samples coming from the head, upper, center, and lower body. Following our previous work, Fig. 7 summarizes the results for the \( k \)-means method implemented in Matlab\textsuperscript{©} [41]. According to these results, the positive predicted samples represent 79.77\%, while the false negatives are 20.23\%.

In this work, we extend this previous approach by including a new dimension. Instead of exploiting only the traveling time, we also consider the concentration level of nanosensors per vessel segment. According to the distinct probability to find nanosensors per vessel (cf. [16, Fig. 10]), this dimension may also contribute to identify their travelling path. The circuits in the HCS present each one a different set of probabilities, which can later be used as patterns to identify the trajectories.

Fig. 8 illustrates the concentration level curves per circuit after averaging the results of the BVS simulation. These are computed for a total of 1000 nanosensors traveling around the body for 300 s. As the figure shows, the total number of nanosensors oscillates around 50 and 100 according to the segment their traveling through. Furthermore, we can observe nanosensors traveling through the arteries and veins (largest concentration) and those traveling through the capillaries (lowest concentration). As we can see in this figure, each circuit provides a different pattern regarding the evolution of the concentration level.

To label the two-dimensional data, we use the most popular \( k \)-means as the unsupervised method, and the Decision Trees as the supervised one due to their low-complex mechanism. This requisite directs critical to conceiving low-power electronics for implantable devices where access to power sources is limited. In future research, we will implement more robust ML methods like deep learning, investigating a balance between performance and power consumption.

A. Introducing the new dimension: Setting up the dataset

The different curves shown in Fig. 8 provide patterns to further identify the traveled circuit. When a nanosensor travels through a circuit, it records the number of neighboring nanosensors on its path every 1 s. Therefore, we assume...
that nanosensors are capable of communicating with close neighbors, e.g., via ultrasound [42] or terahertz [26]. However, nanosensors do not have to communicate with all others, neither to process received messages, as they only need to detect other sensors within their close neighborhood to get an estimate of the concentration level.

By doing so, we implicitly assume that nanosensors are capable of communicating with neighbors, e.g., via ultrasound [42] or terahertz [26]. Upon delivering this concentration record to the Gateway (as a sequence) located at the right Heart (cf. Fig. 1), we can measure its similarity via cross-correlation with the curves depicted in Fig. 8. Considering the maximum of all resulting correlations then provides us with the new metric.\(^4\)

In this way, the Gateway is able to render two-dimensional data as depicted in Fig. 9 (illustrated for the center body). As can be seen from the figure, the sets are more distinguishable than in the case of time dimension only (depicted in [16, Fig. 7 c)). Using this new dimension, the overlap of samples is reduced when including the new dimension provided by the cross-correlator. This data is then used by the unsupervised ML model to identify the different clusters in accordance to each different circuit.\(^5\) The next two subsections describe the implementation and provide partial results of both ML models using this data.

As a remark, although we obtained the curves in Fig. 8 from previous identification of circuits using BVS, in practice they can also be generated according to the dimension of vessels in the human body. Using reported physiological parameters [27], the concentration of blood can be derived per vessel segment, and a similar proportion for the flowing nanosensors in blood can be used.

\(^4\)The cross-correlation procedure is implemented as the optimal receivers in digital transmissions [43].

\(^5\)The data provided by BVS is accessible at https://github.com/jorge-torresgomez/BVS_data, where we also document this dataset through a datasheet [18].

B. Estimating the total of nanosensors per loop: Unsupervised ML module

This module estimates the total of nanosensors per closed loop according to the data provided by the nanosensors to the Gateway, as depicted in Fig. 9. For the ease of implementation, we use the \textit{k-means} method employing functions provided by Matlab [44]. This method aims finding the best partition of samples into clusters in such a way that on each cluster each sample is close to each other as much as possible, and far from samples in other clusters as well. The function implements this partition by looking for the samples that minimize the sum of distances to the cluster’s centroid [45].

We configured the function using a maximum of 4000 iterations and the city block distance. We use this distance in contrast to the Euclidean distance as it emphasizes differences on the two dimensions separately, thus resulting in a larger separation between clusters. The resulting clusters are depicted in Fig. 10. The \textit{k-means} method can successfully identify the different circuits for the travelling nanosensors. In spite of the overlapping sets (e.g. Spleen, Kidneys, and Intestine), the clusters present a distribution close to the original one in Fig. 9.

C. Providing early alerts: Supervised ML module

The supervised module is built using the training data set provided by the unsupervised model. Similar to our previous work in [16], we use a Decision Trees algorithm to predict the location of the abnormality as it provides high accuracy with a low-complexity mechanism [46]. Using Gini’s diversity index as the splitting rule, the resulting supervised method minimizes the errors using the resubstitution estimate [47]. This algorithm is implemented with 100 splits using Gini’s diversity index as the split criterion achieving 85.91 % accuracy. Fig. 11 depicts the positive predicted and false negative samples after integrating the unsupervised (\textit{k-means}) and the supervised (Decision Trees) ML models. In contrast to Fig. 7, the use
VI. RESULTS AND DISCUSSION

We illustrate the results concerning the predicted location of nanosensors for each different vessel segment as well as the results for the detection capabilities of the nanosensors. According to the distribution of nanosensors in the vessels, detection probabilities of abnormalities will be better in some regions compared to others (i.e., regions with the highest distribution have the highest probability of detection).

To compute the stationary distribution of nanosensors per vessel segment we use the Markov model formulation for evaluating Eq. (3). To evaluate it, the transition probabilities (cf. Eq. (5)) are computed according to the estimated number of nanosensors by the unsupervised ML method (cf. Section V-B). In this way, we follow the steps 1, 2 and 3 from Fig. 6 according to the right branch of the flow diagram in Fig. 3.

Fig. 12 depicts the resulting stationary distribution of nanosensors using this methodology after collecting a total of $10^4$ samples at the Gateway. As expected, the most significant probability of finding a nanosensor is in those sinks vessel segments or where all the circuits intercept. For instance, circuits intercept in the arteries as A1, A2, A20 (corresponding to the Arcus aorta, Aorta thoratica, and the Aorta abdominalis), and in the organs and capillaries concerning the heart and the lungs, where the concentration is high. It also happens in the veins as V3 and V5 (Superior and Inferior vena Cava) as blood collectors of the superior and inferior body regions. Although the results presented here are similar to our previous work in [16, Fig. 10], a closer look at the present case exhibits a larger distribution in the Inferior Vena Cava (V5) than the Superior vena Cava (V3). This provides a more realistic description as the Inferior Vena Cava is the sink segment for the center and lower body accounting for more blood flow than the Superior Vena Cava.

To account for the detection capabilities of the flowing nanosensors, we evaluate the expression in Eq. (2). The probability to find a nanosensor in a given tissue ($P_{c,k}$) is directly obtained from Fig. 12, while the probability for a nanosensor to visit a given sensing region (cf. Eq. (1)) is obtained via COMSOL simulations according to the left branch in Fig. 3. Fig. 13 depicts the probability of successful detection versus the total number of nanosensors. By evaluating Eq. (2), this probability is calculated for shoulders and upper arms but can be easily extended to other parts as well. As expected, we observe that the probability of detection increases with increasing number of nanosensors. As the threshold value, defined by $N_{th}$ in Eq. (2), increases, so does the total number of nanosensors required to achieve the same detection probability. Compared to our previous results in [16, Fig. 11], the discrepancy between the curves corresponding to shoulder and arms is reduced. This is due to the fact that the distribution probability in Fig. 12 for shoulders is now closer to the one for upper arms.

In Fig. 14, the probability of successful detection versus the total number of nanosensors is plotted this time for knees and feet. Since the distribution probabilities for both are lower than the ones for shoulders and upper arms, more nanosensors are required to reach the same probability of successful detection. Similar to Fig. 13, with increasing threshold value, the total number of nanosensors required to reach the same probability of detection also increases. Due to the larger volume and the smaller probability of nanosensors being in the feet than in the knees, the number of required nanosensors to detect an infection in the feet is higher than in the knees, as depicted in Fig. 12.

A. Remarks on potential validation

The methodology presented in this paper can be potentially validated with organ-on-chip technology and the current
development of sensors and transceivers in nanoscale size. The research community is devising the complex structures of vessels and organs on chips with high fidelity [48]. Emulating the human body, bio-printed vessels-on-chip like aortas and 3D-capillary networks allow the study of drug pharmacokinetics in fluid environments [49]. Besides, tiny electronic devices, including sensors, processors, memories, and transceivers, allow conceiving nanosensors to sense, actuate and report while flowing [8]. Using these technologies, equipped nanosensors inserted in printed vessel networks may provide a platform to test their detection capabilities while reporting data to external bio-implantable devices [50]. These more powerful external devices may allow running the needed ML methods to identify the traveling paths.

VII. Conclusion

In this work, we investigated a ML-based approach to detect abnormalities in the human body by deploying nanosensors flowing in the bloodstream and sensing their environment. Using the data reported by nanosensors, we evaluated the effectiveness of their detection and localization capabilities of abnormalities in the body. Although we discussed one potential scenario to detect QS molecules released by bacteria, this approach can be extended to detect and localize other chemical components according to the nanosensor capabilities. To that end, we conceived a methodology assembling several parts: i) determining the probability of sensing for nanosensors eavesdropping to quorum sensing communication of infectious bacteria simulated using COMSOL; ii) modeling the traveling path of nanosensors through a Markov chain; iii) computing the transition probabilities using machine learning models. Particularly, in this study, we incorporated a new dimension into our localization dataset for the ML algorithms, i.e., the concentration of neighboring nanosensors perceived by each nanosensor throughout its journey on top of the existing dimension of traveling times of nanosensors. This helped us to improve estimating the locations visited by nanosensors and
to provide a more accurate probability of distribution in the body.

By equipping the nanosensors with different sensing elements, our approach can be applied to biomarker sensing in general, towards building a network of nanodevices enabling in-body precision medicine. We plan to better take the physical activities of humans into account. We will evaluate the performance of the proposed methodology by incorporating different activities, e.g., walking, running, and sleeping. We will also explore better sensing and localization mechanisms tailored for in-body sensing applications.

ACKNOWLEDGMENT

Reported research was supported in part by the project MAMOKO funded by the German Federal Ministry of Education and Research (BMBF) under grant number 16KIS0917 as well as the project NaBoCom funded by the German Research Foundation (DFG) under grant number DR 639/21-1.

REFERENCES


Anke Kuestner received her B.Sc. degree in Computer Science from Paderborn University, Germany in 2020. She has been with the Telecommunication Networks Group, Department of Telecommunication Systems, Technical University of Berlin as a student research assistant from 2020 until 2022. Currently, she is doing her M.Sc. at the Technical University of Berlin.

Jennifer Simonjan received her Ph.D. degree in Information and Communication Engineering from the University of Klagenfurt, Austria in 2019. After her Ph.D. she gathered experience working as a post doc at the Georgia Institute of Technology in Atlanta, US and as a researcher in Austria. Currently, she works as a senior researcher in the Autonomous Robotics Research Center at the Technology Innovation Institute in Abu Dhabi, UAE. Her research interest comprise self-organization and communication in nanoscale and macroscale sensor networks as well as robotic networks. She serves as IEEE PR officer of the Austrian section and as member of several conference organizing committees.

Bige Deniz Unluturk is an assistant professor in both Department of Electrical & Computer Engineering and Biomedical Engineering at Michigan State University since Fall 2021. Between August 2020 - August 2021 she was a Postdoctoral Research Associate in Dr. Chris Contag’s Lab at the Institute for Quantitative Health Science and Engineering, MSU. Before joining MSU, she received her Ph.D. degree in Electrical and Computer Engineering from the Georgia Institute of Technology, Atlanta, GA, in August 2020 under the guidance of Prof. Dr. Ian F. Akyildiz. In 2013, she received her M.Sc. degree from Koc University, Istanbul, Turkey. In 2011 she graduated from Electrical and Electronics Engineering at Middle East Technical University, Ankara, Turkey. Her research is in the areas of wireless communication and networking, more specifically Molecular Communications and Internet of Bio NanoThings and their applications to healthcare.

Falko Dressler is full professor and Chair for Data Communications and Networking at the School of Electrical Engineering and Computer Science, TU Berlin. He received his M.Sc. and Ph.D. degrees from the Dept. of Computer Science, University of Erlangen in 1998 and 2003, respectively. Dr. Dressler has been associate editor-in-chief for IEEE Trans. on Mobile Computing and Elsevier Computer Communications as well as an editor for journals such as IEEE/ACM Trans. on Networking, IEEE Trans. on Network Science and Engineering, Elsevier Ad Hoc Networks, and Elsevier Nano Communication Networks. He has been chairing conferences such as IEEE INFOCOM, ACM MobiSys, ACM MobiHoc, IEEE VNC, IEEE GLOBECOM. He authored the textbooks Self-Organization in Sensor and Actor Networks published by Wiley & Sons and Vehicular Networking published by Cambridge University Press. He has been an IEEE Distinguished Lecturer as well as an ACM Distinguished Speaker. Dr. Dressler is an IEEE Fellow as well as an ACM Distinguished Member. He is a member of the German National Academy of Science and Engineering (acatech). He has been serving on the IEEE COMSOC Conference Council and the ACM SIGMOBILE Executive Committee. His research objectives include adaptive wireless networking (sub-6GHz, mmWave, visible light, molecular communication) and wireless-based sensing with applications in ad hoc and sensor networks, the Internet of Things, and Cyber-Physical Systems.

Jorge Torres Gómez received the B.Sc., M.Sc., and Ph.D. degrees from the Technological University of Havana, CUJAE, Cuba, in 2008, 2010, and 2015, respectively. He is currently with the Telecommunication Networks Group, Department of Telecommunication Systems, Technical University of Berlin. He is holding positions as head of educational activities at the IEEE Germany Section supporting teaching activities and chairing conference events in education. Since 2008, he has been with the School of Telecommunications and Electronics, CUJAE University, where he was a Lecturer with CUJAE, from 2008 to 2018. He has been with the Department of Signal Theory and Communications, Carlos III University of Madrid, Leganés Campus, Madrid, Spain, as a guest lecturer, and with the Department of Digital Signal Processing and Circuit Technology, Chemnitz University of Technology as a postdoc. His research interests include molecular communications, nanoworkets, the age of information, and digital signal processing.