Localization of a Passive Molecular Transmitter with a Sensor Network

Fatih Gulec(✉) and Baris Atakan

Izmir Institute of Technology, 35430 Urla, Izmir, Turkey
{fatihgulec,barisatakan}@iyte.edu.tr

Abstract. Macroscale molecular communication (MC), which has a potential for practical applications, is a promising area for communication engineering. In a practical scenario such as monitoring air pollutants released from an unknown source, it is essential to estimate the location of the molecular transmitter (TX). This paper presents a novel Sensor Network-based Localization Algorithm (SNCLA) for passive transmission by using a novel experimental platform which mainly comprises a clustered sensor network (SN) with 24 sensor nodes and evaporating ethanol molecules as the passive TX. With the usage of the SN concept, novel methods can be developed for the problems in macroscale MC by utilizing the wide literature of sensor networks. In SNCLA, Gaussian plume model is employed to derive the location estimator. The parameters such as transmitted mass, wind velocity, detection time and actual concentration are calculated or estimated from the measured signals via the SN to be employed as the input for the location estimator. The numerical results show that the performance of SNCLA is better for stronger winds in the medium. Our findings show that evaporated molecules do not propagate homogeneously through the SN due to the presence of the wind. In addition, the estimation error of SNCLA decreases for higher detection threshold values.

Keywords: Macroscale molecular communications · Sensor networks · Localization

1 Introduction

Molecular communication (MC) which employs chemical signals for information transfer is an emerging area in the last century for communication engineers [1, 11, 33]. The motivation of MC is that electromagnetic wave-based communication is not suitable for some environments such as human body in microscale [3] or infrastructures comprising pipes in macroscale [15, 38]. The applicability of MC systems in such environments makes the MC a prominent communication paradigm for future applications.

Supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant 119E041.

Y. Chen et al. (Eds.): BICT 2020, LNICST 329, pp. 317–335, 2020.
https://doi.org/10.1007/978-3-030-57115-3_28

In the literature, experimental platforms are proposed for macroscale MC. The first pioneering experimental platform consists of a transmitter (TX) which emits alcohol molecules with a fan behind it, and a receiver (RX) which is an alcohol sensor [8]. The data rate performance of the MC system in [8] is increased by employing multiple transmitters and receivers, i.e., multiple input multiple output (MIMO) technique [22,23]. In [10], an experimental platform which encodes the information symbols by the pH level of the emitted chemicals is proposed. It resembles the human cardiovascular system by using peristaltic pumps as the TX. [41] proposes a similar system given in [10] by using magnetic nanoparticles instead of chemicals. Another experimental platform is accomplished by using an odor generator as the TX and a mass spectrometer as the RX [12,27,28]. Furthermore, MC is proposed to be used in mobile robot platforms [42,45]. In [45], an algorithm is proposed for mobile RX robots to move towards a static sprayer, similar to a bacteria swarm.

The studies employing these experimental platforms are mostly about channel modeling [9,21,24] or just showing that information transfer can be accomplished. However, the estimation of channel parameters such as the number (or mass) of molecules, velocity of the flow in the medium and the distance between the TX and RX is generally not studied for practical scenarios. Especially, the estimation of the distance is important, since the channel parameters can be configured for more efficient communication by estimating the distance accurately [2,7,34]. Furthermore, the distance estimation can also be employed to locate a molecular source in an environmental monitoring application or a virus source propagating through breath in the air [20]. Most of the distance estimation methods which can be classified as two-way and one-way methods are proposed in microscale. In two-way methods, the RX sends a feedback signal to the RX after receiving a pilot signal sent by the TX [29–31]. In one-way methods, the RX estimates the distance from the transmitted signal of the TX [19,25,35,43,44]. All these distance estimation methods are based on the diffusion of molecules in an ideal microscale channel. Only the study in [13] proposes distance estimation methods, which are based on data analysis and machine learning, for a practical macroscale scenario by using experimentally obtained data. It is hard to determine the exact location of the TX with only the distance information. Therefore, localization methods are needed for multi-dimensional practical scenarios. In the literature, a localization algorithm is proposed by using a mobile search robot as the RX moving towards the source according to molecule concentration gradient for a long range underwater scenario [37]. However, the performance of this algorithm is not known for a practical scenario.

All of the platforms given above focus on active transmission of molecules such as sprayers or pumps. However, there is not any platform to understand the dynamics of macroscale MC with passive transmission such as evaporating toxic molecules from a threatening source through the air. Moreover, there is not any experimentally validated localization method for practical macroscale scenarios. Within this context, we propose a novel experimental platform for macroscale MC applications and a novel localization algorithm by using this platform.

Firstly, our experimental platform consists of a passive source which include freely evaporating ethanol molecules. The experimental platform is placed in a fume hood which is a closed box to provide controlled conditions. Evaporating molecules are detected by a sensor network (SN) which includes 24 MQ-3 alcohol sensor nodes in a rectangular order. The novelty of our experimental platform lies in the usage of the SN which can pave the way to novel methods by adapting techniques from the SN literature. Moreover, the concept of employing a SN can be applied for different practical scenarios such as the localization of an underwater molecular TX.

Secondly, the Sensor Network-based Clustered Localization Algorithm (SNCLA) is proposed for the localization of a passive molecular TX as a proof-of-concept application employing our experimental platform. The SN is divided into four clusters. Primarily, the Gaussian plume model which is employed widely in the meteorology literature to model the movement of the pollutant particles in the air is given as the system model. As for the SNCLA, the location estimator is derived for the sensor node pairs in each cluster. In order to use the location estimator, some experimental parameters such as the actual concentration, transmitted mass and the wind velocity are required to be estimated or calculated. To this end, the measured sensor voltage is smoothed by using a moving average filter and a detection is made according to a predetermined detection threshold voltage. The measured sensor voltages at the chosen threshold voltage are converted to actual concentration values via the sensitivity response of the sensors. The detection time of the SN is employed to estimate the velocity of the wind in the medium for four directions on the $x - y$ plane. The estimated wind velocity is taken as the input for the mass calculation of the evaporated molecules. The location estimator employs all these estimated/calculated values as the input. Finally, SNCLA determines two clusters according to the magnitude of the wind velocities estimated for the four directions and makes the location estimation for the sensor nodes in these clusters. The numerical results show that SNCLA performs better, when the wind velocity is higher. Furthermore, the average detection times for all of the sensor nodes are given to show the propagation pattern of the evaporating molecules. Surprisingly, the evaporating molecules do not propagate in an isotropic fashion. It is observed that there is always a wind in the medium that affects the propagation of molecules. In addition, cluster error is defined as an error metric to evaluate the performance of the clusters in the SN. It is shown that cluster errors decrease and more stable results can be obtained for higher detection thresholds.

The remainder of the paper is organized as follows. In Sect. 2, the experimental platform is given in detail. Section 3 introduces the system model on which the SNCLA is based. The SNCLA is presented in Sect. 4. The numerical results are shown and analyzed in Sect. 5. Finally, the concluding remarks are given in Sect. 6.

2 Experimental Platform

In this section, the experimental platform which is employed for the localization of a molecular transmitter using a SN is introduced. As shown in Fig. 1, this platform consists of a TX and a SN placed inside a fume hood, which is a closed cabinet to conduct chemical experiments at controlled conditions without being exposed to chemicals. The TX includes a pipette pump, two pipettes, a rubber hose and a circular petri dish. The pipette connected to the tip of the pipette pump is filled with liquid ethanol before the transmission. When liquid ethanol is pumped through the rubber hose, it fills the petri dish which has a radius of 2.25 cm. The petri dish is deployed at the midpoint of the SN. The transmission is realized by the evaporation of ethanol molecules in the petri dish at room temperature (25 °C). After the transmission, evaporated ethanol molecules propagate in the air.

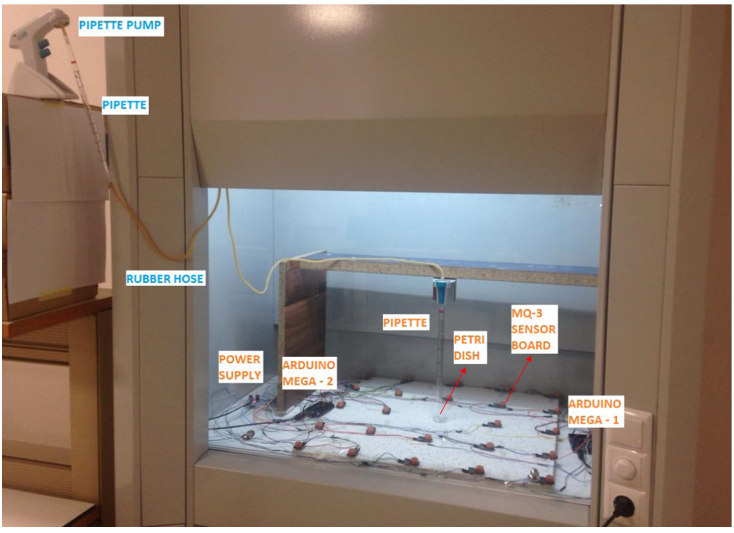

Fig. 1. Experimental platform.

The SN consists of 24 MQ-3 alcohol sensor boards (or nodes), a power supply and two Arduino Mega microcontroller boards which are connected to a computer. The sensor nodes are placed on a rectangular surface of 60×60 cm. The distance between two adjacent nodes on the horizontal and vertical axis is 15 cm. Each sensor board has a 1 kΩ load resistor on it in order to generate a voltage to be an analog input signal for the microcontroller board. As shown in Fig. 1, while fourteen of the sensor nodes are wired to the first Arduino microcontroller board, ten of them are connected to the second Arduino board. In order to synchronize the TX and SN, the nodes start to receive signals as soon as the petri dish is filled with 5 ml of liquid ethanol. Next, the system model

to explain the propagation of evaporated molecules employed in the proposed experimental platform is given.

3 System Model

This section details the system model on which the localization algorithm is based. In macroscale MC, diffusion-based models are employed to explain the propagation of molecules through the air [9,27,45]. Unlike macroscale experimental studies in the literature, molecules are released by evaporation at room temperature in our scenario. Since there is not any applied force for the emission, we can classify it as a passive transmission. The absence of the force applied to this emission makes molecules susceptible to the effects of wind or flows in the air, even at low velocities. Actually, there is almost always a slight wind in the air [17]. Considering the passive transmission of molecules and winds in the air, Gaussian plume model, which is widely used in the meteorology literature for the dispersion of air pollutants, can be applied for our scenario to model the propagation of evaporated molecules. By using the conservation of mass, the equation below can be written as [40]

$$\frac{\partial C}{\partial t} + \nabla \cdot \vec{J} = S, \tag{1}$$

where S is the source term, $\vec{J}$ and C represent the mass flux and concentration of evaporated molecules, respectively. Here, the mass flux can be given as the summation of diffusive flux $\vec{J_D}$, which stems from the turbulent diffusivity in the atmosphere, and the advective flux $\vec{J_a}$ stemming from the wind velocity ($\vec{u}$). Hence, the mass flux is given as

$$\vec{J} = \vec{J_D} + \vec{J_a} = -\vec{K}\nabla C + C\vec{u}, \tag{2}$$

where $\vec{K} = diag(K_x, K_y, K_z)$ is a diagonal matrix showing the turbulent diffusivities in three dimensions. Thus, (1) takes the form of the equation which is known as the atmospheric diffusion (or dispersion) equation as given by [40]

$$\frac{\partial C}{\partial t} + \nabla \cdot (C\vec{u}) = \nabla \cdot (\vec{K}\nabla C) + S. \tag{3}$$

In our scenario, we define the TX at the position (x_T, y_T, z_T) as an instantaneous source to have a time-dependent solution in (3). In fact, our experimental platform is in a sufficiently small scale so that the TX can be considered as a source releasing molecules in an instantaneous puff. Furthermore, the wind velocity is defined with two components in x and y axes, i.e., u_x and u_y. It is assumed that the plane at $z = 0$ is a reflective plane and there is not any other boundaries. For the source term which is defined as $S = \frac{m_T}{\vec{u}}\delta(x-x_T)\delta(y-y_T)\delta(z-z_T)\delta(t)$ where m_T is the transmitted mass and $\delta(.)$ is the Dirac delta function, the solution of (3) is given as [6]

$$C(x,y,z,t) = \frac{(\pi t)^{-3/2} m_T}{8(K_x K_y K_z)^{1/2}} \exp\left(-\frac{(x-x_T-u_x t)^2}{4K_x t} - \frac{(y-y_T-u_y t)^2}{4K_y t}\right) \times \left[\exp\left(-\frac{(z-z_T)^2}{4K_z t}\right) + \exp\left(-\frac{(z+z_T)^2}{4K_z t}\right)\right], \quad (4)$$

which is known as the Gaussian puff solution. Here, $e^{-\frac{(z+z_T)^2}{4K_z t}}$ represents the reflection of the plume from the ground. In the literature of atmospheric dispersion, the turbulent diffusivities are defined in terms of dispersion parameters such that $\sigma_x^2 = 2K_x t$, $\sigma_y^2 = 2K_y t$, $\sigma_z^2 = 2K_z t$ [39]. Hence, (4) is rearranged as

$$C(x,y,z,t) = \frac{m_T}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} \exp\left(-\frac{(x-x_T-u_x t)^2}{2\sigma_x^2} - \frac{(y-y_T-u_y t)^2}{2\sigma_y^2}\right) \times \left[\exp\left(-\frac{(z-z_T)^2}{2\sigma_z^2}\right) + \exp\left(-\frac{(z+z_T)^2}{2\sigma_z^2}\right)\right]. \quad (5)$$

The advantage of this conversion is to determine the dispersion parameters (σ_x, σ_y, σ_z) by using empirically derived models which depend on the distance between the TX and RX. According to the model given in [5] which is widely used in the meteorology literature, σ_y and σ_z for stable air conditions as in our case are calculated by

$$\sigma_y = \frac{0.04r}{(1+0.0001r)^{0.5}} \quad (6)$$

$$\sigma_z = \frac{0.016r}{(1+0.0003r)} \quad (7)$$

where r is the distance to the source in meters and σ_x can be approximated as $\sigma_x \approx \sigma_y$ [6]. Regarding these empirical models in (6) and (7), the effect of the dispersion parameters on the system model are negligible, since the scale of our SN is small (60×60 cm). Therefore, the dispersion parameters are defined as constant values. In our scenario, the SN and TX are all deployed at $z = 0$. Accordingly for each sensor, the concentration is given as

$$C_{i,j} = \frac{m_T}{\sqrt{2\pi^3}\sigma_x\sigma_y\sigma_z} \exp\left(-\frac{(x_{i,j}-x_T-u_x t_{i,j})^2}{2\sigma_x^2} - \frac{(y_{i,j}-y_T-u_y t_{i,j})^2}{2\sigma_y^2}\right), \quad (8)$$

where $(x_{i,j}, y_{i,j})$ and $t_{i,j}$ show the location and detection time of the node $N_{i,j}$ which is in the i^{th} row and j^{th} column of the SN, respectively. Here, $i = 1, ..., M_r$ and $j = 1, ..., M_c$.

In addition, the deployment of the sensor nodes are illustrated in Fig. 2. It is assumed that each sensor node knows its location in the Cartesian coordinate system. As shown in this figure, the SN is divided into four clusters. These clusters are employed for the localization algorithm of the TX in MC as given in the next section.

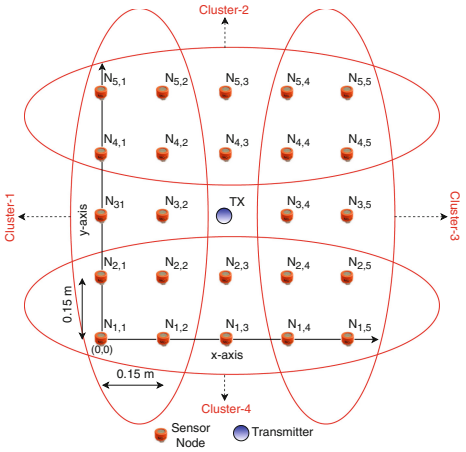

Fig. 2. The deployment of the sensor nodes and TX.

4 Sensor Network-Based Clustered Localization Algorithm

In this section, Sensor Network-Based Clustered Localization Algorithm (SNCLA) whose block diagram is given in Fig. 3 is proposed. First, the location estimator is derived using the system model given in Sect. 3. Then, the estimation and calculation of the required parameters for the location estimator is detailed. At the end of this section, SNCLA is detailed by employing all the estimated and calculated parameters.

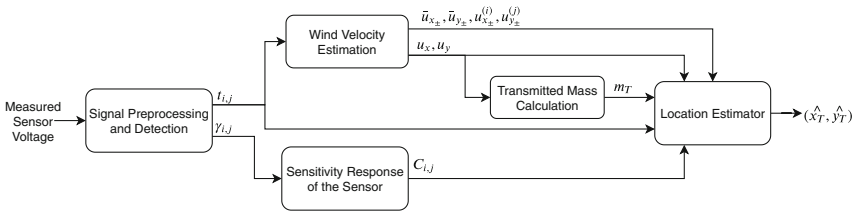

Fig. 3. Block diagram of the SNCLA.

4.1 Derivation of the Location Estimator

In order to derive the location estimator, (8) can be written as

$$\frac{\sqrt{2\pi^3}\sigma_x\sigma_y\sigma_z C_{i,j}}{m_T} = \exp\left(-\frac{(x_{i,j}-x_T-u_x t_{i,j})^2}{2\sigma_x^2} - \frac{(y_{i,j}-y_T-u_y t_{i,j})^2}{2\sigma_y^2}\right). \quad (9)$$

When the natural logarithm, i.e., ln(.), of both sides is taken, then (9) is given by

$$\ln\left(\frac{\sqrt{2\pi^3}\sigma_x\sigma_y\sigma_z C_{i,j}}{m_T}\right) = -\frac{(x_{i,j}-x_T-u_x t_{i,j})^2}{2\sigma_x^2} - \frac{(y_{i,j}-y_T-u_y t_{i,j})^2}{2\sigma_y^2}. \quad (10)$$

For convenience, let $n_{i,j} = \ln\left(\frac{\sqrt{2}(\pi)^{3/2}\sigma_x\sigma_y\sigma_z C_{i,j}}{m_T}\right)$. Hence, the final equation for the location estimator is given by

$$\frac{(x_{i,j}-x_T-u_x t_{i,j})^2}{2\sigma_x^2} + \frac{(y_{i,j}-y_T-u_y t_{i,j})^2}{2\sigma_y^2} + n_{i,j} = 0. \quad (11)$$

For two sensor nodes, a system of nonlinear equations can be generated using (11) where x_T and y_T are the variables and the other parameters are constant. Since the solution of this system is not easily tractable, numerical methods can be used to obtain the solution as detailed later in this section. In order to solve these equations, parameters such as $C_{i,j}$, $t_{i,j}$, m_T and wind velocity values are required to be estimated or calculated.

4.2 Signal Preprocessing and Detection

In our experimental platform, the concentration is measured as a voltage value from the sensor nodes. Due to the random movements of molecules, there are fluctuations on the measured sensor voltage. In order to detect the signals more accurately, the received signal by the sensor is needed to be smoothed via removing the fluctuations. Therefore, a moving average filter is employed as defined by [36]

$$y[n] = \frac{1}{L}\sum_{k=0}^{L} C[n-k]. \quad (12)$$

where $C[n]$ is the measured sensor voltage, $y[n]$ and L are the output and window size of the filter, respectively.

When there is no transmission from the TX, the sensors still output a positive voltage value which is defined as the offset level, i.e., $A_{o_{i,j}}$. Since the offset levels of the sensors can be different, the threshold voltage ($\gamma_{i,j}$) is defined for each sensor node by employing a constant detection threshold amplitude (A_T) as given by

$$\gamma_{i,j} = A_{o_{i,j}} + A_T. \quad (13)$$

In order to calculate $A_{o_{i,j}}$, the first p samples of the received signal is averaged before the moving average filter. After $\gamma_{i,j}$ is determined for each sensor node, the time instances that each sensor reaches the $\gamma_{i,j}$ value in $y[n]$ is recorded as $t_{i,j}$. These $t_{i,j}$ and $\gamma_{i,j}$ values are employed as input for velocity estimation and the sensitivity response of the sensor, respectively.

4.3 Sensitivity Response of the Sensor

Sensor voltage is obtained via a sensor measurement circuit on the MQ-3 sensor boards which is shown in Fig. 4. As the concentration around the sensor changes, its resistance (R_s) changes. Hence, the molecule concentration is converted to an electrical signal via the circuit in Fig. 4 where V_{out} gives the output voltage. Using this circuit, R_s is derived as

$$R_s = \left(\frac{V_{in}}{V_{out}} - 1\right) R_l, \tag{14}$$

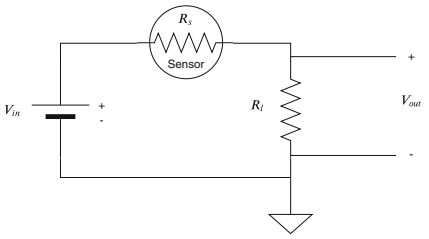

Fig. 4. Measurement circuit of the sensor board.

where V_{in} shows the DC input voltage and R_l is the load resistance. For each concentration value, the sensor has a different R_s value. The sensor resistance can be normalized by dividing R_s to R_o where R_o is the sensor resistance measured at the concentration value of $0.0004\,\mathrm{kg/m^3}$ which is the minimum concentration level MQ-3 sensor can measure [18]. According to its datasheet, MQ-3 sensor has a sensitivity characteristic which maps each concentration value to the normalized resistance value (R_s/R_o) [18]. This sensitivity characteristic can be expressed as a sensitivity function

$$f\left(C_{i,j}\right) = \frac{R_s}{R_o} = \left(\frac{V_{in}}{V_{out}} - 1\right) \frac{R_l}{R_o}, \tag{15}$$

where $C_{i,j}$ is the actual molecule concentration around the sensor and R_s/R_o is given by substituting (14) into (15). By employing the values in its datasheet, $f(C_{i,j})$ can be obtained via curve fitting technique. Nonlinear least squares method that minimizes the sum of the square errors is employed to fit the datasheet values of the MQ-3 sensitvity characteristic. As the result of the curve fitting, $f(C_{i,j})$ is given by

$$f\left(C_{i,j}\right) = a_1 \left(C_{i,j}\right)^{b_1} + d_1, \tag{16}$$

where a_1, b_1 and d_1 are the curve fitting parameters. By employing Levenberg-Marquardt algorithm, these parameters are estimated as $a_1 = 0.0116$,

$b = -0.5855$ and $d_1 = -0.0743$ with a Root Mean Square Error (RMSE) value of 0.0371 [16]. The sensitivity response of the MQ-3 sensor is also employed in our study [14] which is a part of the signal reconstruction approach of the RX in macroscale. The signal reconstruction of the RX is first proposed in [4] in order to investigate how the actual concentration around the RX is sensed in microscale.

In order to find the molecule concentration for the given detection threshold voltage, V_{out} is set as $\gamma_{i,j}$ in (15). (15) and (16) are combined to obtain the equation as given by

$$\left(\frac{V_{in}}{V_{out}} - 1\right)\frac{R_l}{R_o} = a_1 \left(C_{i,j}\right)^{b_1} + d_1. \tag{17}$$

As the result of the sensitivity response of the sensor, (17) is manipulated to obtain $C_{i,j}$ which is given by

$$C_{i,j} = \left(\frac{V_{in}R_l - \gamma_{i,j}R_l - d_1\gamma_{i,j}R_o}{\gamma_{i,j}R_o a_1}\right)^{(1/b_1)}. \tag{18}$$

4.4 Wind Velocity Estimation

After the threshold voltages ($\gamma_{i,j}$) and detection times ($t_{i,j}$) are obtained for each sensor, the wind velocity flowing over two sensor nodes can be estimated in x and y directions as given by

$$u_x = \frac{|x_2 - x_1|}{t_2 - t_1}, u_y = \frac{|y_2 - y_1|}{t_2 - t_1}, \tag{19}$$

where (x_1, y_1) and (x_2, y_2) are the coordinates for the first and second sensor node, respectively and t_1 and t_2 are the detection times for the first and second sensor node, respectively. For our scenario, (19) is generalized by averaging the wind velocities estimated by the sensor node pairs in the corresponding cluster defined in Fig. 3 for four directions according to the formulas given below

$$\bar{u}_{x_-} = \frac{1}{M_r}\sum_{i=1}^{M_r} u_{x_-}^{(i)} = \frac{1}{M_r}\sum_{i=1}^{M_r}\frac{|x_{i,1} - x_{i,2}|}{(t_{i,1} - t_{i,2})}, \text{ Cluster-1 (–x direction)} \tag{20}$$

$$\bar{u}_{x_+} = \frac{1}{M_r}\sum_{i=1}^{M_r} u_{x_+}^{(i)} = \frac{1}{M_r}\sum_{i=1}^{M_r}\frac{|x_{i,5} - x_{i,4}|}{(t_{i,5} - t_{i,4})}, \text{ Cluster-3 (+x direction)} \tag{21}$$

$$\bar{u}_{y_+} = \frac{1}{M_c}\sum_{j=1}^{M_c} u_{y_+}^{(j)} = \frac{1}{M_c}\sum_{j=1}^{M_c}\frac{|y_{5,j} - y_{4,j}|}{(t_{5,j} - t_{4,j})}, \text{ Cluster-2 (+y direction)} \tag{22}$$

$$\bar{u}_{y_-} = \frac{1}{M_c}\sum_{j=1}^{M_c} u_{y_-}^{(j)} = \frac{1}{M_c}\sum_{j=1}^{M_c}\frac{|y_{1,j} - y_{2,j}|}{(t_{1,j} - t_{2,j})}, \text{ Cluster-4 (–y direction)} \tag{23}$$

where $u_{x_\pm}^{(i)}$ and $u_{y_\pm}^{(j)}$ show the instantaneous wind velocity of the i^{th} and j^{th} sensor node pair in the corresponding direction (or cluster), respectively, $\bar{u}_{x_\pm}$ and $\bar{u}_{y_\pm}$ show the average of these sensor node pair velocities in the corresponding direction (or cluster), respectively, $x_{i,j}$ and $y_{i,j}$ indicate the horizontal and vertical position of the sensor node $N_{i,j}$ given in Fig. 2, respectively and M_r and M_c are the total number of rows and columns of the SN, respectively. Here, the instantaneous velocities whose values are negative are not considered for the velocity estimation. During the experiments it is observed that the wind blows stronger in one direction which means that it can only have at most two velocity components among the estimated velocities in four directions. Therefore, u_x and u_y are defined as:

$$u_x = \max(\bar{u}_{x_-}, \bar{u}_{x_+}), \quad u_y = \max(\bar{u}_{y_-}, \bar{u}_{y_+}). \tag{24}$$

4.5 Transmitted Mass Calculation

The estimated values of u_x and u_y are used to calculate the evaporation rate of ethanol in the air (Q_e). For the wind blowing over a surface with a velocity u at room temperature (25 °C), Q_e (kg/m^{2}s) is given by [26]

$$Q_e = h_1 u^{0.54}, \tag{25}$$

where $h_1 = 4 \times 10^{-3}$ kg/m^3 and $u = \sqrt{u_x^2 + u_y^2}$. In order to find the mass flow rate of evaporated molecules, i.e., Q (kg/s), which is defined as the mass flowing through a surface per unit time, $Q = Q_e A$ where A is the surface area of evaporated molecules, i.e., the surface area of the petri dish for our case. Here, the instantaneous puff of the TX which is represented by $\delta(t)$ in the system model is approximated by a pulse with a short emission time (T_e). Hence, the transmitted mass can be calculated as [32]

$$m_T = QT_e = Q_e A T_e. \tag{26}$$

4.6 Operation of the SNCLA

Thus far, the required input parameters for the location estimator (see Fig. 3) are obtained. By using these parameters, Algorithm 1 is proposed for the localization of the TX. In this algorithm, two clusters are chosen according to the direction of the wind velocity on x and y axes. For instance, if the wind blows stronger in the $+x$ direction on the x-axis and $+y$ direction on the y axis, then the node pairs in Cluster-3 and Cluster-2 are chosen for the location estimation. Similar to the wind velocity estimation, the node pairs whose absolute instantaneous wind velocity value is negative are not considered for the location estimation. Afterwards, the equation pairs given in (27)–(30) are solved for x_T and y_T according to the chosen two clusters. The solution for each of two equations gives the estimated coordinates of the TX, i.e., $\hat{x}_T$ and $\hat{y}_T$. The solutions of

Algorithm 1. SNCLA

input: $\bar{u}_{x_\pm}$, $\bar{u}_{y_\pm}$, $u^{(i)}_{x_\pm}$, $u^{(j)}_{y_\pm}$, m_T, $C_{i,j}$, $t_{i,j}$ for all $i = 1, ..., M_r$, $j = 1, ..., M_c$
if $(u_x == u_{x_-})$ and $(u_y == u_{y_+})$ **then**
    Calculate $(\hat{x_T}, \hat{y_T})$ by (27) for Cluster 1
    Calculate $(\hat{x_T}, \hat{y_T})$ by (28) for Cluster 2
else if $(u_x == u_{x_-})$ and $(u_{y_-} == u_{y_+})$ **then**
    Calculate $(\hat{x_T}, \hat{y_T})$ by (27) for Cluster 1
    Calculate $(\hat{x_T}, \hat{y_T})$ by (30) for Cluster 4
else if $(u_x == u_{x_+})$ and $(u_y == u_{y_+})$ **then**
    Calculate $(\hat{x_T}, \hat{y_T})$ by (28) for Cluster 2
    Calculate $(\hat{x_T}, \hat{y_T})$ by (29) for Cluster 3
else
    Calculate $(\hat{x_T}, \hat{y_T})$ by (29) for Cluster 3
    Calculate $(\hat{x_T}, \hat{y_T})$ by (30) for Cluster 4
end if

(27)–(30) are too long to write in this paper. Instead, these equations are solved numerically as given in the numerical results.

$$\left.\begin{aligned} \frac{(x_{i,1} - x_T - u_x t_{i,1})^2}{2\sigma_x^2} + \frac{(y_{i,1} - y_T - u_y t_{i,1})^2}{2\sigma_y^2} + n_{i,1} = 0 \\ \frac{(x_{i,2} - x_T - u_y t_{i,2})^2}{2\sigma_x^2} + \frac{(y_{i,2} - y_T - u_y t_{i,2})^2}{2\sigma_y^2} + n_{i,2} = 0 \end{aligned}\right\} \; i = 1, ..., M_r. \quad \text{(Cluster-1)} \tag{27}$$

$$\left.\begin{aligned} \frac{(x_{5,j} - x_T - u_x t_{5,j})^2}{2\sigma_x^2} + \frac{(y_{5,j} - y_T - u_y t_{5,j})^2}{2\sigma_y^2} + n_{5,j} = 0 \\ \frac{(x_{4,j} - x_T - u_x t_{4,j})^2}{2\sigma_x^2} + \frac{(y_{4,j} - y_T - u_y t_{4,j})^2}{2\sigma_y^2} + n_{4,j} = 0 \end{aligned}\right\} \; j = 1, ..., M_c. \quad \text{(Cluster-2)} \tag{28}$$

$$\left.\begin{aligned} \frac{(x_{i,5} - x_T - u_x t_{i,5})^2}{2\sigma_x^2} + \frac{(y_{i,5} - y_T - u_y t_{i,5})^2}{2\sigma_y^2} + n_{i,5} = 0 \\ \frac{(x_{i,4} - x_T - u_y t_{i,4})^2}{2\sigma_x^2} + \frac{(y_{i,4} - y_T - u_y t_{i,4})^2}{2\sigma_y^2} + n_{i,4} = 0 \end{aligned}\right\} \; i = 1, ..., M_r. \quad \text{(Cluster-3)} \tag{29}$$

$$\left.\begin{aligned} \frac{(x_{1,j} - x_T - u_x t_{1,j})^2}{2\sigma_x^2} + \frac{(y_{1,j} - y_T - u_y t_{1,j})^2}{2\sigma_y^2} + n_{1,j} = 0 \\ \frac{(x_{2,j} - x_T - u_x t_{2,j})^2}{2\sigma_x^2} + \frac{(y_{2,j} - y_T - u_y t_{2,j})^2}{2\sigma_y^2} + n_{2,j} = 0 \end{aligned}\right\} \; j = 1, ..., M_c. \quad \text{(Cluster-4)} \tag{30}$$

5 Numerical Results

In this section, numerical results of the SNCLA is given. 25 measurements each lasting 180 s were performed with the experimental platform. There were at least 30 min left among adjacent measurements in order to decrease the concentration

Table 1. Experimental parameters.

Parameter	Value
Number of measurements (M_m)	25
Detection threshold amplitude (A_T)	0.055 V
Emission time (T_e)	0.1 s
Actual TX location (x_T, y_T)	(0.3, 0.3) m
Area of the petri dish (A)	0.0024 m^2
Input voltage of the sensor board (V_{in})	5 V
Load resistance (R_l)	1 kΩ
Sensor resistance at 0.0004 kg/m^3 (R_o)	24 kΩ
Standard deviation of the Gaussian concentration distribution on the x, y and z axis ($\sigma_x, \sigma_y, \sigma_z$)	0.0115 m, 0.0115 m, 0.0046 m
Number of samples to be averaged for the offset level $A_{o_{i,j}}$ of the sensors (p)	3
Window size of the moving average filter (L)	7

level with the ventilation of the fume hood. The ventilation was not used during the measurements.

The experimental parameters are given in Table 1. Among these parameters, R_o is calculated by employing output voltage (V_{out}) of the sensor measurement circuit. According to the MQ-3 sensor datasheet, its detection scope is between 5×10^{-5} and 10^{-2} kg/m^3 [18]. This detection scope is scaled for V_{out} values between 0 and 5 V. Thus, 0.0004 kg/m^3 corresponds to $V_{out} = 0.2$ V which is used to calculate the sensor resistance value via (14). As mentioned in Sect. 3, the dispersion parameters ($\sigma_x, \sigma_y, \sigma_z$) are assumed as constant values. For our experimental scenario and according to (6)–(7), the ranges of σ_y and σ_z are 0.006–0.017 m and 0.0024–0.0068 m. Therefore, σ_y and σ_z are chosen as the average values of these ranges. σ_x is also taken as equal to σ_y [6]. The window size of the moving average filter (L) and number of samples to be averaged to determine the offset level ($A_{o_{i,j}}$) of the sensor node $N_{i,j}$, i.e., p, is determined empirically for our experimental scenario. In addition, the detection threshold amplitude (A_T) is also chosen as an empirical value in order to have accurate estimations. However, the error values for a range of A_T values are also given in Fig. 8.

The estimation results for each cluster are shown in Fig. 5. For the experimental values, $\hat{x}_T$ and $\hat{y}_T$ have two complex conjugate roots for each. Therefore, only real parts of the solutions are considered for the numerical results. As shown in Fig. 5, there are more results for Cluster-1 and Cluster-2, since the wind velocity is mostly in $-x$ and $+y$ direction for our measurements. The results of these two clusters are also more accurate than the other clusters, since the wind blows stronger in $-x$ and $+y$ direction than the other directions. The accuracy of these

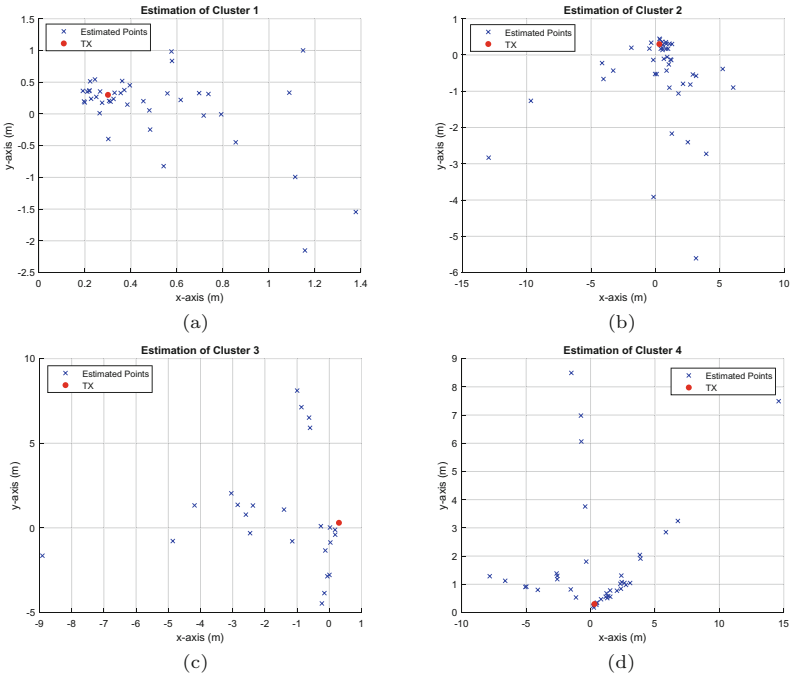

Fig. 5. Estimated points using SNCLA for each cluster.

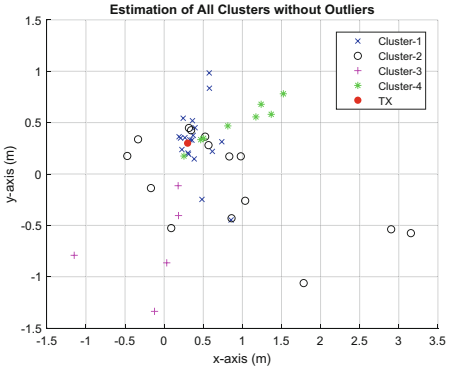

Fig. 6. Estimated points using SNCLA for all clusters without outliers.

clusters are more clearly depicted in Fig. 6. This figure shows the results of the best ten measurements in the same scale for a better visual perception of the figure. When the wind velocity is higher, SNCLA gives better results, since the effect of the dispersion of evaporated molecules decreases.

In Fig. 7, the average of the detection times for each sensor node is given as a heatmap. This figure verifies the direction of the wind by the detection times.

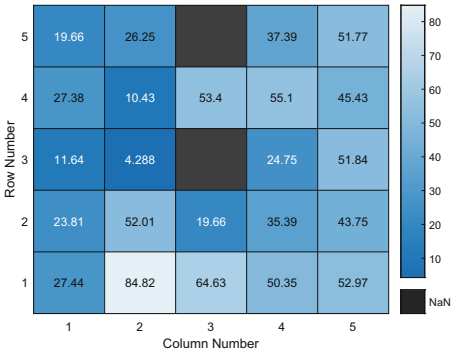

Fig. 7. Average of the detection times for each sensor node.

Interestingly, there is no detection for the given threshold for the sensor node $N_{5,3}$. Actually, there are also very few detections for the other sensor nodes in the third column of the SN. According to these observations, some of the evaporated molecules move in the same direction with the wind whereas the rest of the molecules move mostly in the opposite horizontal direction of the wind due to the initial puff of the TX.

For the last part of the numerical results, an error metric, which is called Cluster Error (ϵ_c), is defined for each cluster by

$$\epsilon_c = \begin{cases} \frac{1}{M_r}\sum_{i=1}^{M_r}\frac{1}{M_m}\sum_{k=1}^{M_m}\sqrt{(x_T-\hat{x}_{T_k}^{(i)})^2+(y_T-\hat{y}_{T_k}^{(i)})^2}, & \text{Cluster 1,3} \quad (31) \\ \frac{1}{M_c}\sum_{j=1}^{M_c}\frac{1}{M_m}\sum_{k=1}^{M_m}\sqrt{(x_T-\hat{x}_{T_k}^{(j)})^2+(y_T-\hat{y}_{T_k}^{(j)})^2}, & \text{Cluster 2,4} \quad (32) \end{cases}$$

where M_m is the number of the measurements and $(\hat{x}_{T_k}^{(i)}, \hat{y}_{T_k}^{(i)})$ show the estimated points for the i^{th} node pair in the corresponding cluster at the k^{th} measurement. In our case, there are $M_r = M_c = 5$ node pairs for each cluster. First, the Euclidean distance between the actual and estimated points are calculated for each node pair in the cluster. Then, these distances for all the measurements are averaged. This process is repeated for each node pair in the cluster. The results of ϵ_c for A_T values between 0–0.15 V with 0.001 V steps are given in Fig. 8. For higher threshold values, Cluster-1 and 2 outperform the other clusters due to the higher wind velocities. For lower thresholds, Cluster-3 has better results due to the lower number of detections. Figure 8 also shows that the choice of the detection threshold for lower error is significant which is left as an open research issue.

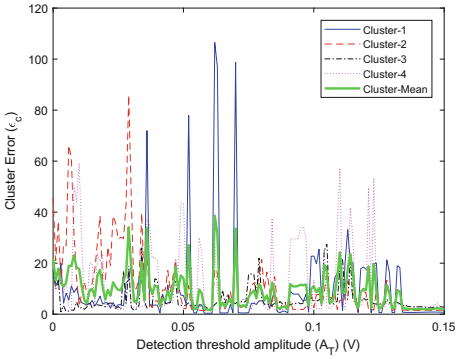

Fig. 8. Cluster errors for different detection threshold amplitudes.

6 Conclusion

This paper presents a novel experimental platform for macroscale MC and a novel algorithm for the localization of a molecular TX with a sensor network of four clusters, i.e., SNCLA. In our experimental platform, the molecular TX emits molecules by evaporation at room temperature and the signals are received with the SN. First, Gaussian plume model is given as the system model for our scenario. Based on this system model, a location estimator is derived. Then, estimation/calculation methods for the unknown parameters in the location estimator such as detection time, transmitted mass, wind velocity and the actual concentration are proposed. Finally, SNCLA is explained by combining all these estimated/calculated parameters. In SNCLA, the estimated location of the TX is based on the estimated wind velocity direction and the derived location estimator. SNCLA gives more accurate results for the clusters in the same direction with the wind for higher detection threshold values. Since the Gaussian plume model on which the SNCLA is based is employed for longer distances in the meteorology domain, it is anticipated to have more accurate results on larger scales with the proposed SNCLA. As the future work, we plan to improve this model on larger scales and adapt different localization algorithms from the sensor network literature with MC perspective.

References

1. Atakan, B.: Molecular Communications and Nanonetworks. Springer, New York (2014). https://doi.org/10.1007/978-1-4939-0739-7
2. Atakan, B., Akan, O.B.: An information theoretical approach for molecular communication. In: Bio-Inspired Models of Network, Information and Computing Systems, Bionetics 2007, 2nd edn., pp. 33–40. IEEE (2007)
3. Atakan, B., Akan, O.B., Balasubramaniam, S.: Body area nanonetworks with molecular communications in nanomedicine. IEEE Commun. Mag. **50**(1), 28–34 (2012)

4. Atakan, B., Gulec, F.: Signal reconstruction in diffusion-based molecular communication. Trans. Emerg. Telecommun. Technol. **30**(12), e3699 (2019)
5. Briggs, G.A.: Diffusion estimation for small emissions. Atmospheric turbulence and diffusion laboratory, p. 83 (1973)
6. De Visscher, A.: Air Dispersion Modeling: Foundations and Applications. Wiley, Hoboken (2013)
7. Eckford, A.W.: Achievable information rates for molecular communication with distinct molecules. In: Bio-Inspired Models of Network, Information and Computing Systems, Bionetics 2007, 2nd edn, pp. 313–315. IEEE (2007)
8. Farsad, N., Guo, W., Eckford, A.W.: Tabletop molecular communication: text messages through chemical signals. PLoS ONE **8**(12), e82935 (2013)
9. Farsad, N., Kim, N.R., Eckford, A.W., Chae, C.B.: Channel and noise models for nonlinear molecular communication systems. IEEE J. Sel. Areas Commun. **32**(12), 2392–2401 (2014)
10. Farsad, N., Pan, D., Goldsmith, A.: A novel experimental platform for in-vessel multi-chemical molecular communications. In: GLOBECOM 2017–2017 IEEE Global Communications Conference, pp. 1–6. IEEE (2017)
11. Farsad, N., Yilmaz, H.B., Eckford, A., Chae, C.B., Guo, W.: A comprehensive survey of recent advancements in molecular communication. IEEE Commun. Surv. Tutor. **18**(3), 1887–1919 (2016)
12. Giannoukos, S., Marshall, A., Taylor, S., Smith, J.: Molecular communication over gas stream channels using portable mass spectrometry. J. Am. Soc. Mass Spectrom. **28**(11), 2371–2383 (2017)
13. Gulec, F., Atakan, B.: Distance estimation methods for a practical macroscale molecular communication system. Nano Commun. Netw. **24**, 100300 (2020)
14. Gulec, F., Atakan, B.: A fluid dynamics approach to channel modeling in macroscale molecular communication. arXiv preprint arXiv:2004.03321 (2020)
15. Guo, W., Mias, C., Farsad, N., Wu, J.L.: Molecular versus electromagnetic wave propagation loss in macro-scale environments. IEEE Trans. Mol. Biol. Multi-Scale Commun. **1**(1), 18–25 (2015)
16. Hagan, M.T., Menhaj, M.B.: Training feedforward networks with the marquardt algorithm. IEEE Trans. Neural Netw. **5**(6), 989–993 (1994)
17. Hanna, S.R., Briggs, G.A., Hosker Jr., R.P.: Handbook on atmospheric diffusion. Technical report, National Oceanic and Atmospheric Administration, Oak Ridge, TN, USA (1982)
18. Hanwei Electronics Co., Ltd.: Technical data of MQ-3 gas sensor (2018)
19. Huang, J.T., Lai, H.Y., Lee, Y.C., Lee, C.H., Yeh, P.C.: Distance estimation in concentration-based molecular communications. In: Global Communications Conference (GLOBECOM), pp. 2587–2591. IEEE (2013)
20. Khalid, M., Amin, O., Ahmed, S., Shihada, B., Alouini, M.S.: Communication through breath: aerosol transmission. IEEE Commun. Mag. **57**(2), 33–39 (2019)
21. Kim, N.R., Farsad, N., Chae, C.B., Eckford, A.W.: A universal channel model for molecular communication systems with metal-oxide detectors. In: 2015 IEEE International Conference on Communications (ICC), pp. 1054–1059. IEEE (2015)
22. Koo, B.H., Lee, C., Yilmaz, H.B., Farsad, N., Eckford, A., Chae, C.B.: Molecular mimo: from theory to prototype. IEEE J. Sel. Areas Commun. **34**(3), 600–614 (2016)
23. Lee, C., et al.: Molecular mimo communication link. In: 2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 13–14. IEEE (2015)

24. Lee, C., Yilmaz, H.B., Chae, C.B., Farsad, N., Goldsmith, A.: Machine learning based channel modeling for molecular mimo communications. arXiv preprint arXiv:1704.00870 (2017)
25. Lin, L., Luo, Z., Huang, L., Luo, C., Wu, Q., Yan, H.: High-accuracy distance estimation for molecular communication systems via diffusion. Nano Commun. Netw. **19**, 47–53 (2019)
26. Lyulin, Y.V., Feoktistov, D.V., Afanasev, I.A., Chachilo, E.S., Kabov, O.A., Kuznetsov, G.V.: Measuring the rate of local evaporation from the liquid surface under the action of gas flow. Tech. Phys. Lett. **41**(7), 665–667 (2015). https://doi.org/10.1134/S1063785015070251
27. McGuiness, D.T., Giannoukos, S., Marshall, A., Taylor, S.: Parameter analysis in macro-scale molecular communications using advection-diffusion. IEEE Access **6**, 46706–46717 (2018)
28. McGuiness, D.T., Giannoukos, S., Taylor, S., Marshall, A.: Experimental and analytical analysis of macro-scale molecular communications within closed boundaries. IEEE Trans. Mol. Biol. Multi-Scale Commun. **5**, 44–55 (2019)
29. Moore, M., Nakano, T., Enomoto, A., Suda, T.: Measuring distance with molecular communication feedback protocols. In: Proceedings of ICST BIONETICS, pp. 1–13 (2010)
30. Moore, M.J., Nakano, T.: Comparing transmission, propagation, and receiving options for nanomachines to measure distance by molecular communication. In: 2012 IEEE International Conference on Communications (ICC), pp. 6132–6136. IEEE (2012)
31. Moore, M.J., Nakano, T., Enomoto, A., Suda, T.: Measuring distance from single spike feedback signals in molecular communication. IEEE Trans. Signal Process. **60**(7), 3576–3587 (2012)
32. Munson, B.R., Young, D.F., Okiishi, T.H., Huebsch, W.W.: Fundamentals of Fluid Mechanics. Wiley, Hoboken (2009)
33. Nakano, T., Eckford, A.W., Haraguchi, T.: Molecular Communication. Cambridge University Press, Cambridge (2013)
34. Nakano, T., Okaie, Y., Vasilakos, A.V.: Transmission rate control for molecular communication among biological nanomachines. IEEE J. Sel. Areas Commun. **31**(12), 835–846 (2013)
35. Noel, A., Cheung, K.C., Schober, R.: Joint channel parameter estimation via diffusive molecular communication. IEEE Trans. Mol. Biol. Multi-Scale Commun. **1**(1), 4–17 (2015)
36. Oppenheim, A.V.: Discrete-Time Signal Processing. Pearson Education India, Delhi (1999)
37. Qiu, S., et al.: Long range and long duration underwater localization using molecular messaging. IEEE Trans. Mol. Biol. Multi-Scale Commun. **1**(4), 363–370 (2015)
38. Qiu, S., Guo, W., Wang, S., Farsad, N., Eckford, A.: A molecular communication link for monitoring in confined environments. In: 2014 IEEE International Conference on Communications Workshops (ICC), pp. 718–723. IEEE (2014)
39. Seinfeld, J.H., Pandis, S.N.: Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. Wiley, Hoboken (2016)
40. Stockie, J.M.: The mathematics of atmospheric dispersion modeling. Siam Rev. **53**(2), 349–372 (2011)
41. Unterweger, H., et al.: Experimental molecular communication testbed based on magnetic nanoparticles in duct flow. In: 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 1–5. IEEE (2018)

42. Wang, L., Farsad, N., Guo, W., Magierowski, S., Eckford, A.W.: Molecular barcodes: information transmission via persistent chemical tags. In: 2015 IEEE International Conference on Communications (ICC), pp. 1097–1102. IEEE (2015)
43. Wang, X., Higgins, M.D., Leeson, M.S.: An algorithmic distance estimation scheme for diffusion based molecular communication systems. In: 2015 IEEE International Conference on Communications (ICC), pp. 1134–1139. IEEE (2015)
44. Wang, X., Higgins, M.D., Leeson, M.S.: Distance estimation schemes for diffusion based molecular communication systems. IEEE Commun. Lett. **19**(3), 399–402 (2015)
45. Zhai, H., Liu, Q., Vasilakos, A.V., Yang, K.: Anti-ISI demodulation scheme and its experiment-based evaluation for diffusion-based molecular communication. IEEE Trans. Nanobiosci. **17**(2), 126–133 (2018)