Breath-Source Localization in Air-Based Molecular Communication: A Learning-Driven Proof-of-Concept

Sunasheer Bhattacharjee, Saswati Pal, Péter Scheepers, and Falko Dressler
School of Electrical Engineering and Computer Science, TU Berlin, Germany
{bhattacharjee,pal,scheepers,dressler}@ccs-labs.org

Abstract

Exhaled breath is a biologically generated molecular signal exhibiting spatiotemporal humidity–temperature dynamics. Beyond respiratory pattern recognition, interpreting breath as an information-bearing signal enables non-invasive monitoring of airway function and supports future wearable and ambient health-sensing applications. This work studies mouth-nose source localization using a single low-cost DHT22 sensor. Temporal data from Eupnea, Bradypnea, and Tachypnea are processed using learning-driven methods for source inference under a subject-independent evaluation protocol, achieving a mean balanced accuracy of 76% across leave-one-subject-out folds, with up to 87% on an individual held-out subject. The results demonstrate that even minimal sensing hardware captures discriminative spatiotemporal features while ensuring generalization across unseen subjects, establishing a proof-of-concept for breath-source localization in air-based molecular communication.

CCS Concepts

• **Hardware → Sensors and actuators**; • **Applied computing → Health care information systems**.

Keywords

IoBNT, Localization, Molecular Communication, Exhaled Breath, Sensing, Machine Learning, Information Theory

ACM Reference Format:
Sunasheer Bhattacharjee, Saswati Pal, Péter Scheepers, and Falko Dressler. 2026. Breath-Source Localization in Air-Based Molecular Communication: A Learning-Driven Proof-of-Concept. In *13th Annual ACM International Conference on Nanoscale Computing and Communication (NANOCOM '26), September 21–23, 2026, St. John's, NL, Canada.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3818305.3830250

1 Introduction

The Internet of Bio-Nano Things targets nanoscale biological and artificial devices communicating via molecular or terahertz signals [1, 7], but clinical deployment is limited by miniaturization, energy, and biocompatibility. As a practical alternative, the Internet of Bio Things exploits biological transmitters, primarily humans, and engineered receivers through air-based molecular communication (ABMC) [2, 13], enabling non-invasive breath monitoring and linking molecular communication (MC) theory to deployable healthcare.

ABMC-based breath sensing aims for low-cost, unobtrusive respiratory monitoring beyond clinical settings. Exhaled breath is modeled as a molecular signal whose source and dynamics can be inferred using simple receivers.

Unlike conventional clinical breath diagnostics conducted under controlled supervision, this work targets passive and ambient sensing scenarios where the exhalation origin cannot be assumed *a priori*, as individuals switch between nasal and oral breathing depending on physiology, activity, or airway condition. In such settings, user input is impractical, motivating automatic source localization.

Exhaled breath encodes airway, metabolic, and inflammatory states through gases and aerosols including carbon dioxide (CO_2), humidity, and volatile organic compounds (VOCs) [17]. Building on our prior work, which showed that humidity-temperature signatures from a single DHT22 sensor can classify respiratory patterns via machine learning (ML) [3], we extend ABMC to mouth-nose localization. Although both pathways carry the same biomarkers, their origin is diagnostically distinct: oral nitric oxide (NO) elevation indicates eosinophilic asthma [6], reduced nasal NO suggests primary ciliary dyskinesia [21], and oral versus nasal volatile profiles differentiate periodontal disease from sinonasal inflammation [12, 18]. The oral-versus-nasal breathing ratio is further associated with nasal obstruction, sleep-disordered breathing, and respiratory fatigue.

This work establishes a minimal-hardware sensing-and-decoding framework for breath-source localization within an ABMC perspective. Modeling mouth–nose as distinct molecular sources, we investigate whether a single low-cost DHT22 sensor can support mouth–nose localization through supervised learning under a subject-independent leave-one-subject-out (LOSO) evaluation protocol. As an initial proof-of-concept, our objective is to demonstrate the feasibility of breath-source localization using a low-cost humidity–temperature sensor.

Our key contributions can be summarized as follows:

- Breath-source localization formulated as a decoding problem within ABMC
- Minimal-hardware, wearable-compatible experimental setup based on humidity–temperature temporal sensing
- Learning-based subject-independent localization framework with information-theoretic characterization

2 Related Work

Breath-based sensing has been widely explored for non-invasive health monitoring, especially for detecting gaseous biomarkers such as CO_2, NO, and VOCs. While prior studies show strong diagnostic value for respiratory and metabolic disorders, they commonly

NANOCOM '26, St. John's, NL, Canada

ACM ISBN 979-8-4007-2767-2/2026/09
https://doi.org/10.1145/3818305.3830250

model exhalation as a spatially uniform source, neglecting anatomical origin [9, 14, 17]. Within ABMC, exhaled breath is modeled as an information-bearing signal propagating through an ambient channel [2, 13], and recent studies show that simple receivers can classify respiratory patterns using temporal features [3]. These works assume a single biological source of emission and do not address source localization, thereby neglecting spatial ambiguity in biomarker release. Clinical and biomechanical studies have compared nasal and oral breathing using spirometry [8], imaging [22], or multi-sensor setups to characterize airflow [20]. While informative, these methods rely on bulky instruments, limiting low-cost sensing and communication-theoretic modeling, and they do not leverage propagation-induced signal dynamics for source inference.

Recent ML approaches demonstrate that temporal features from low-cost sensors can classify breathing patterns, but they often aggregate signals across pathways or assume fixed sources [10, 11]. As a result, diagnostically important cases, where identical biomarkers from the mouth or nose indicate different pathologies, remain largely unaddressed. Localization has been studied in MC, e.g., ML-based infection source localization in biological channels [16]. However, existing works focus on estimation or classification performance rather than information-theoretic analysis.

We address these limitations by introducing breath-source localization within ABMC, modeling mouth-nose as two distinct and separable molecular sources in space. Using a single low-cost humidity–temperature sensor and supervised learning on a combined dataset of Eupnea, Bradypnea, and Tachypnea, we investigate the feasibility of mouth-nose discrimination under a minimal sensing setup. The approach is evaluated under a subject-independent framework, highlighting its behavior across individuals and the influence of subject-specific variability. This formulation additionally provides an initial information-theoretic perspective by treating the anatomical source as an information-bearing symbol observed through a noisy ambient channel, thereby bridging ABMC theory with wearable-compatible diagnostics.

3 System Model

Figure 1 depicts the proposed ABMC system, where exhaled breath is modeled as a molecular signal emitted by a biological source, propagating through air, and detected by a low-cost receiver. The objective is to localize the emitting source, mouth or nose, from the received signal. The model is not intended to provide an accurate physical description of breath propagation, which would require detailed fluid-dynamic modeling of turbulent airflow. Instead, it serves as a conceptual abstraction that explains how humidity and temperature signatures originating from mouth and nose exhalations are transformed into sensor measurements and motivates the subsequent data-driven classification framework. The model describes the response to a single exhalation event. Temporal pulse broadening is inherently accounted for through the channel and sensor impulse responses.

3.1 Source Model

The mouth $\mathcal{M}$ and nose $\mathcal{N}$ are modeled as discrete molecular sources. The random variable $X \in \{\mathcal{M}, \mathcal{N}\}$ denotes which source produced a given exhalation event, and each event is assumed to originate from exactly one source. During an exhalation from source X, air carrying water vapor molecules is released into the environment. The water vapor concentration of the exhaled air at the source boundary is modeled as

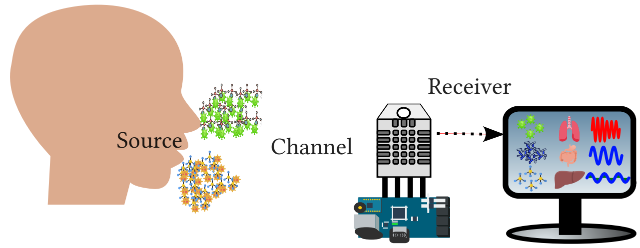


Figure 1: System model for analyzing the exhaled breath from a source via a channel to a receiver.

$$n_X(t) = c_X\, s_X(t), \tag{1}$$

where c_X is the peak water vapor concentration of the exhaled air, and $s_X(t)$ is a normalized temporal waveform that captures the airflow envelope of the exhalation event.

3.2 Channel Model

After emission, the exhaled plume disperses through the ambient air via advection, turbulent mixing, and molecular diffusion. Over the short observation window and source–receiver distances considered, this dispersal and mixing process is approximated as a linear time-invariant system. Under this approximation, the water vapor concentration at the receiver location is

$$n_{X,\text{channel}}(t) = n_X(t) * h(t), \tag{2}$$

where $h(t)$ is the effective impulse response that captures the spatiotemporal dilution of the plume between the source and the receiver.

3.3 Receiver Model

A single DHT22 sensor at the receiver location measures humidity and temperature, both of which are perturbed by the arriving exhaled plume. The sensor is approximated as a first-order low-pass system with impulse response $g(t)$. The receiver observation model is

$$\mathbf{y}_X(t) \triangleq \begin{bmatrix} y_{X,\text{H}}(t) \\ y_{X,\text{T}}(t) \end{bmatrix} = \mathbf{f}\big(n_{X,\text{channel}}(t) * g(t)\big) + \mathbf{w}(t), \tag{3}$$

where $\mathbf{f}(\cdot)$ maps the local water vapor concentration to the humidity and temperature levels observed by the sensor, and $\mathbf{w}(t)$ is an additive noise process. Since the exhaled plume co-transports heat and moisture, the local water vapor concentration serves as a sufficient proxy for both humidity and temperature perturbations at the sensor. The sensor is also assumed to be linear time-invariant system over the observation window. As the mouth and nose exhalations produce different spatiotemporal plume characteristics, the resulting time series signal at the receiver exhibits source-dependent signatures. Each received signal is mapped to features from which

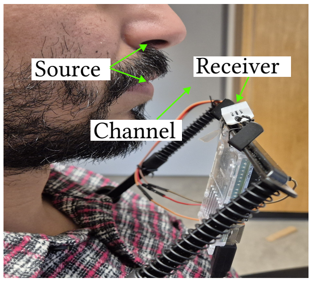


Figure 2: Sensor assembly as used in our experiments. The sensor is positioned close to mouth and nose.

the source estimate $\hat{X} \in \{\mathcal{M}, \mathcal{N}\}$ is inferred. These systematic differences motivate the use of feature-based supervised ML models to infer the source identity from the measured signals.

4 Experimental Setup

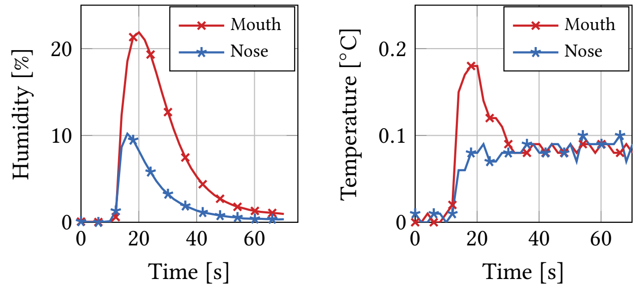
4.1 Hardware Configuration

A DHT22 humidity–temperature sensor interfaced with an Arduino Uno recorded exhaled breath during Eupnea, Bradypnea, and Tachypnea at 2 s sampling intervals. This sampling rate was chosen considering the intrinsic response time of the DHT22 sensor, which is on the order of several seconds. Consequently, the measurements capture only slower plume dynamics rather than rapid intra-breath transients. Mounted on a harmonica stand and positioned ∼5 cm in front of the nostrils, the sensor maintained line-of-sight to peak nasal airflow while limiting oral-flow bias (cf. Figure 2). Experiments were conducted indoors under stable ambient conditions, with the setup ensuring consistent placement across trials. The resulting setup should therefore be viewed as a controlled laboratory prototype for feasibility evaluation rather than as a deployable system. In practical applications, the sensing unit could be miniaturized and integrated into wearable or smart-environment platforms for unobtrusive respiratory monitoring.

4.2 Data Collection

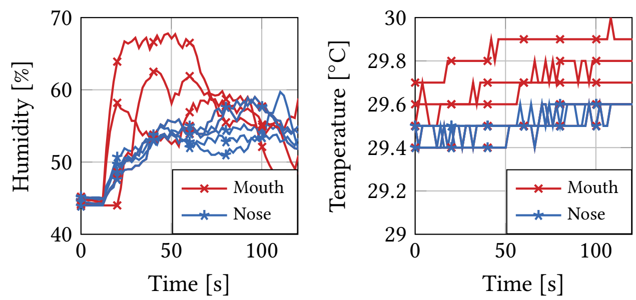
Experiments were conducted with three healthy adult subjects (two male, one female) aged 25–35. None of the subjects reported respiratory or airway-related conditions during data collection. Each subject performed controlled mouth-only and nose-only breathing trials. Eupnea represents natural resting breathing, while Bradypnea and Tachypnea were voluntarily emulated based on standard clinical ranges. Specifically, Eupnea corresponds to 12–20 breaths/min, Bradypnea to below 12 breaths/min, and Tachypnea to above 20 breaths/min [23]. Subjects followed verbal pacing instructions to maintain consistent breathing intervals. Each recording corresponds to a single controlled trial and contains synchronized humidity and temperature time series. In total, 1021 recordings were collected across all subjects, sources, and patterns, including 512 mouth- and 509 nose-breathing samples.

Figure 3 shows the baseline-corrected system impulse response (SIR) for humidity and temperature, averaged over 20 trials to reduce environmental variability. Subjects performed a brief exhalation through either the mouth or nose while maintaining a fixed



(a) Humidity **(b) Temperature**

Figure 3: System impulse response averaged over 20 trials and baseline corrected for both mouth and nose.



(a) Humidity **(b) Temperature**

Figure 4: Humidity and temperature trends for Eupnea over 5 trials each from mouth and nose.

position relative to the sensor. For the SIR characterization only, no additional breaths were taken during the observation window, allowing the impulse response to decay naturally. A 70 s duration was selected to capture both the initial response and subsequent relaxation. As shown in Figure 3a, mouth breathing produces a sharp, high-amplitude humidity impulse due to larger flow and a wider plume, whereas nose breathing yields a weaker, more gradual response. After exhalation, humidity returns to baseline within 60–70 s, governed by diffusion, ambient mixing, and DHT22 recovery. These rise–decay profiles reflect source-dependent airflow characteristics. Similarly, Figure 3b shows that mouth breathing produces a higher temperature offset due to stronger convective heat transfer. Unlike humidity, temperature does not return to baseline within the observation window because of the DHT22's thermal inertia. Heat retention in the sensor and enclosure leads to slow relaxation and a quasi-steady elevated offset.

Figure 4 shows representative Eupnea trials (five each for mouth and nose). From Figure 4a, mouth breathing exhibits a rapid rise to higher humidity with greater inter-trial variability due to stronger, turbulent airflow, whereas nose breathing increases more gradually and remains tightly clustered. The observed drift and plateau reflect sensor response and ambient mixing. Figure 4b further shows that mouth breathing produces a consistent positive temperature offset, while nose breathing remains lower and more stable with minimal spread. Despite some overlap between trials, thermal signatures still provide complementary information for distinguishing breath sources.

5 Learning-based Localization Framework

At the receiver, the goal is to localize the breath source from the humidity–temperature response recorded by the DHT22 sensor. Differences in airflow geometry and exhalation dynamics between mouth and nose breathing lead to distinct temporal signatures in the measured signals, which are summarized through handcrafted descriptors and exploited for classification. Breath-source localization is therefore formulated as a binary supervised classification problem. For trial i, the recorded receiver signal is represented as a bivariate time series

$$\mathbf{R}_i = \begin{bmatrix} r_{1,1} & r_{1,2} \\ r_{2,1} & r_{2,2} \\ \vdots & \vdots \\ r_{\tau_i,1} & r_{\tau_i,2} \end{bmatrix} \in \mathbb{R}^{\tau_i \times 2}, \tag{4}$$

where $r_{t,1}$ and $r_{t,2}$ denote the humidity and temperature measurements at time index t, respectively, and τ_i is the number of samples in trial i. We map each trial to a fixed-dimensional handcrafted feature vector $\mathbf{f}_i \in \mathbb{R}^d$, which is paired with a binary label $y_i \in \{\mathcal{M}, \mathcal{N}\}$, where $y_i = \mathcal{M}$ and $y_i = \mathcal{N}$ correspond to mouth and nose breathing, respectively. Each trial is represented by 76 handcrafted features extracted from the humidity and temperature signals and their first-order differences, capturing distributional statistics, temporal dynamics, signal energy, and cross-channel relationships. These include summary statistics, slope and dominant spectral peak frequency, successive-difference descriptors, zero-crossing behavior, humidity–temperature correlation, mean inter-signal difference and ratio, as well as recording duration and sampling characteristics. After feature extraction, the dataset is given as

$$\mathcal{D} = \{(\mathbf{f}_i, y_i, s_i)\}_{i=1}^{N}, \tag{5}$$

where s_i denotes the subject identity associated with trial i, and N is the total number of trials. To assess generalization across individuals and avoid subject-level information leakage, the dataset is split at the subject level. Since trials were recorded from three subjects, we employ a LOSO evaluation protocol. Although this results in only three evaluation folds, LOSO represents the most conservative assessment possible for the available dataset, as every test fold consists entirely of an unseen individual, thereby preventing subject-level information leakage. For a held-out subject identity u, the LOSO test set is

$$\mathcal{D}_{\text{test}}^{(u)} = \{(\mathbf{f}_i, y_i, s_i) \in \mathcal{D} : s_i = u\}, \tag{6}$$

and the training set consists of the remaining trials

$$\mathcal{D}_{\text{train}}^{(u)} = \{(\mathbf{f}_i, y_i, s_i) \in \mathcal{D} : s_i \neq u\}. \tag{7}$$

This procedure is repeated over all subjects so that each subject is used once as the unseen test subject. Within each LOSO fold, model development is carried out using only the training subjects. All preprocessing operations, including missing-value handling, feature normalization, and decision-threshold selection, are performed exclusively on the training portion of the data and then applied unchanged to the held-out subject. This setup prevents subject-level information leakage and provides a conservative estimate of cross-subject localization performance.

Table 1: LOSO subject-independent localization performance (mean ± SD across held-out subjects).

Model	Accuracy	Bal. Acc.	F1-score	ROC-AUC
CatBoost	**0.76 ± 0.10**	**0.76 ± 0.10**	**0.74 ± 0.12**	**0.83 ± 0.10**
LogReg	0.65 ± 0.11	0.65 ± 0.11	0.64 ± 0.14	0.77 ± 0.12

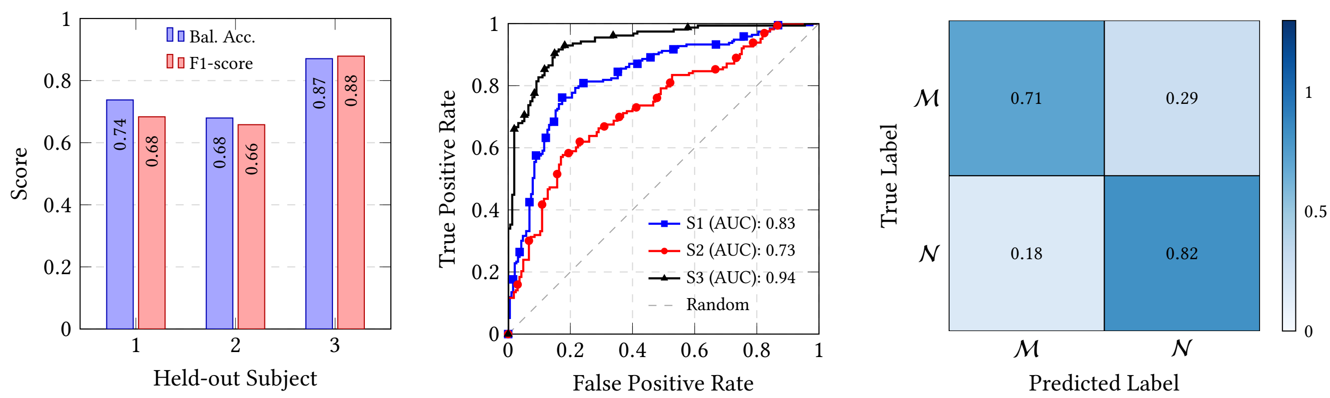
We employ logistic regression [5] and CatBoost [19] to determine whether breath-source localization can be captured by a linear decision rule or requires a more expressive nonlinear model. Logistic regression is used as a linear baseline with median imputation, feature standardization, and L2 regularization (`max_iter=5000`), whereas CatBoost models nonlinear interactions among the handcrafted humidity–temperature features using `iterations=150`, `depth=3`, `learning_rate=0.05`, and `l2_leaf_reg=7` under a Logloss objective. Predicted labels are obtained by thresholding the model posterior output. In each LOSO fold, the decision threshold is selected from the training data by maximizing balanced accuracy on out-of-fold posterior predictions over a grid in $[0.20, 0.80]$. The resulting model is then evaluated on the held-out subject using standard binary-classification metrics.

6 Results

6.1 Performance of the Model

Table 1 shows that subject-independent breath-source localization from humidity–temperature measurements is feasible and that nonlinear modeling provides a clear advantage over a linear baseline. In particular, CatBoost consistently outperforms logistic regression across accuracy, balanced accuracy, F1-score, and receiver operating characteristic area under the curve (ROC-AUC), indicating that the engineered feature space contains useful source-discriminative information that is not adequately captured by a linear classifier. Reporting both accuracy and balanced accuracy is important here as the former reflects overall correctness, whereas the latter indicates whether the two classes are recognized more evenly. The agreement between accuracy and balanced accuracy suggests that the performance gain is not due to a bias toward one class, but reflects a more reliable separation of nose and mouth breathing under subject-independent testing. At the same time, the variability across LOSO folds highlights the challenge of achieving consistent generalization across unseen subjects.

Figure 5 illustrates the CatBoost performance under LOSO evaluation from multiple perspectives. Figure 5a shows that the subject-wise balanced accuracy and F1-score vary across held-out subjects. Balanced accuracy reflects class-balanced recognition, whereas F1-score reflects the precision–recall tradeoff. CatBoost achieved its strongest performance for Subject 3, with balanced accuracy and F1-score of 0.87 and 0.88, respectively, while the lower scores for Subjects 1 and 2 indicate that inter-subject variability remains a key factor in localization performance. The ROC curves in Figure 5b support the same trend, showing stronger source discrimination for Subject 3 and weaker separation for the other held-out subjects.



(a) Mean subject-wise LOSO scores **(b) Subject-wise ROC curves** **(c) Mean normalized confusion matrix**

Figure 5: Performance of the CatBoost model under LOSO evaluation.

The mean normalized confusion matrix in Figure 5c further shows that the classifier retains discriminative capability under subject-independent testing. The model correctly identifies class $\mathcal{M}$ in 71% of cases and class $\mathcal{N}$ in 82% of cases. However, class $\mathcal{M}$ is more frequently misclassified as class $\mathcal{N}$ than vice versa. These results indicate that the proposed humidity-temperature feature representation supports the feasibility of mouth-nose localization under subject-wise evaluation, despite the minimal sensing setup. At the same time, the performance gap across held-out subjects suggests that inter-subject variability in breathing behavior and source–sensor geometry remains the principal factor limiting subject-independent localization. Preliminary CatBoost hyperparameter tuning did not improve LOSO performance; hence, the fixed-parameter CatBoost model is retained. All associated experimental data, script, and output files are publicly available in [4] under CC BY and MIT licenses.

6.2 Information-Theoretic Interpretation

Breath-source localization is formulated as an information transmission problem in which the true breathing source X is inferred from sensor measurements by the trained ML-based classifier. Detection outcomes are assumed conditionally independent across trials given the true source, which justifies this abstraction. The classifier output $Y = \hat{X}$ is modeled as the output of an empirical binary decision channel at the selected operating threshold. From the empirical confusion matrix, the transition probability matrix is constructed as

$$\mathbf{P}_{Y|X} = \begin{bmatrix} P_{\mathcal{M}|\mathcal{M}} & P_{\mathcal{M}|\mathcal{N}} \\ P_{\mathcal{N}|\mathcal{M}} & P_{\mathcal{N}|\mathcal{N}} \end{bmatrix} = \begin{bmatrix} 0.71 & 0.29 \\ 0.18 & 0.82 \end{bmatrix}, \tag{8}$$

where $P_{i|j} \triangleq P(Y = i \mid X = j)$. Since $P_{\mathcal{M}|\mathcal{N}} \neq P_{\mathcal{N}|\mathcal{M}}$, the channel is asymmetric and therefore cannot be reduced to a binary symmetric channel characterized by a single crossover probability.

Based on this model, mutual information (MI) $I(X;Y)$ quantifies the statistical information about the true source retained in the classifier's hard decisions. The MI expression for the binary asymmetric channel [15] for the breath-source localization problem becomes

$$\begin{aligned} I(X;Y) =& H_b\big((1-\alpha)P_{\mathcal{M}|\mathcal{N}} + \alpha(1-P_{\mathcal{N}|\mathcal{M}})\big) \\ &- (1-\alpha)H_b(P_{\mathcal{M}|\mathcal{N}}) - \alpha H_b(P_{\mathcal{N}|\mathcal{M}}), \end{aligned} \tag{9}$$

where $\alpha = P_X(\mathcal{M}) = 1 - P_X(\mathcal{N}) \approx 0.501$ and $H_b(p) = -p\log_2 p - (1-p)\log_2(1-p)$ denotes the binary entropy function. Evaluating (9) yields $I(X;Y) \approx 0.217$ bits per classified recording, a compact summary of the classifier's empirical discriminability rather than an achievable communication rate.

6.3 Discussion

The results support the central premise of this work that mouth-nose localization can be interpreted as a source-decoding problem in macroscopic ABMC. Despite the use of a single low-cost humidity–temperature sensor, the received breath signal retains sufficient structure to support subject-independent inference of source origin. The main limitation is the marked performance variation across held-out subjects, which points to inter-subject and source–sensor geometric variability as the primary barrier to robust generalization. Accordingly, the significance of this work lies in establishing a minimal-hardware proof-of-concept for breath-source localization, while indicating that improved robustness will require stronger invariance to subject-specific breathing and measurement conditions. As the evaluation is based on only three subjects, the reported subject-independent performance should be interpreted as a pilot-scale result rather than a population-level generalization estimate.

To provide a complementary perspective, we model the empirical confusion matrix as a discrete memoryless decision channel. This links the observed classification behavior to a compact information-transfer characterization of the classifier's discriminability. However, as this analysis is derived from a limited dataset with only three subjects, it should be interpreted as an initial characterization rather than a fundamental sensing limit; larger and more diverse data may reveal different levels of information preservation.

From an application perspective, practical deployment will require robustness to variation in source–sensor distance, orientation, user motion, and ambient conditions, as well as to scenarios involving multiple individuals or overlapping airflow sources. Addressing these challenges through improved sensor placement strategies, subject-robust feature design, and multimodal sensing is an important direction for translating this proof-of-concept into a practical breath-monitoring system.

7 Conclusion

We present a minimal ABMC-based framework for breath-source localization, modeling the mouth and nose as distinct molecular transmitters. Using a single low-cost humidity–temperature sensor and ML, we show that spatiotemporal plume dynamics feasibly enable mouth-nose discrimination across core respiratory patterns. The results establish a wearable-compatible baseline and provide a compact information-theoretic characterization of the classifier's discriminability in ABMC. While the present work focuses on controlled measurements with a single sensing node, practical deployment will require addressing environmental variability, sensor placement sensitivity, and hardware response-time limitations.

Future work will address these limitations by expanding the participant cohort, including subjects with respiratory conditions, evaluating robustness under varying sensor placement, source–sensor distance, user motion, and ambient airflow, and investigating faster-response multimodal sensing. Incorporating physics-informed learning grounded in ABMC may further improve robustness, interpretability, and information preservation across subjects.

Acknowledgments

Reported research was supported in part by the project IoBNT, funded by the German Federal Ministry of Research, Technology and Space (BMFTR) under grant number 16KIS1986K, and by the project NaBoCom, funded by the German Research Foundation (DFG) under grant number DR 639/21-2.

References

[1] Ian F. Akyildiz, Massimiliano Pierobon, Sasitharan Balasubramaniam, and Yevgeni Koucheryavy. 2015. The Internet of Bio-Nano Things. *IEEE Communications Magazine* 53, 3 (March 2015), 32–40. doi:10.1109/MCOM.2015.7060516

[2] Sunasheer Bhattacharjee, Dadi Bi, Pit Hofmann, Alexander Wietfeld, Sophie Becke, Michael Lommel, Pengjie Zhou, Ruifeng Zheng, Ulrich Kertzscher, Yansha Deng, Wolfgang Kellerer, Frank H. P. Fitzek, and Falko Dressler. 2026. Exhaled Breath Analysis Through the Lens of Molecular Communication: A Survey. *IEEE Communications Surveys & Tutorials* 28 (Jan. 2026), 412–445. doi:10.1109/COMST.2025.3605748

[3] Sunasheer Bhattacharjee, Saswati Pal, Peter Scheepers, and Falko Dressler. 2025. Breath Patterns as Signals: A Machine Learning-based Molecular Communication Perspective. In *12th ACM International Conference on Nanoscale Computing and Communication (NANOCOM 2025)*. ACM, Chengdu, China, 22–27. doi:10.1145/3760544.3764127

[4] Sunasheer Bhattacharjee, Saswati Pal, Peter Scheepers, and Falko Dressler. 2026. *Breath-Source Localization in Air-Based Molecular Communication: Experimental Data and Reproducibility Package*. Experimental Dataset version 1.0.0. Zenodo. doi:10.5281/zenodo.21378656

[5] D. R. Cox. 1958. The Regression Analysis of Binary Sequences. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 20, 2 (July 1958), 215–232. doi:10.1111/j.2517-6161.1958.tb00292.x

[6] Fernanda Cruvinel de Abreu, Jose Laerte Rodrigues da Silva Júnior, and Marcelo Fouad Rabahi. 2019. The fraction exhaled nitric oxide as a biomarker of asthma control. *Biomarker Insights* 14 (Jan. 2019). doi:10.1177/1177271919826550

[7] Falko Dressler and Stefan Fischer. 2015. Connecting In-Body Nano Communication with Body Area Networks: Challenges and Opportunities of the Internet of Nano Things. *Elsevier Nano Communication Networks* 6 (June 2015), 29–38. doi:10.1016/j.nancom.2015.01.006

[8] R. Gayatri Devi and Gowri Sethu. 2018. Establishing norms for nasal spirometry. *National Journal of Physiology, Pharmacy and Pharmacology* 8, 8 (June 2018), 1188–1193. doi:10.5455/njppp.2018.8.0414226042018

[9] I. Horváth, J. Hunt, and Peter J. Barnes. 2005. Exhaled breath condensate: Methodological recommendations and unresolved questions. *European Respiratory Journal* 26, 3 (Sept. 2005), 523–548. doi:10.1183/09031936.05.00029705

[10] Ilona Karpiel, Maciej Mysiński, Kamil Olesz, and Marek Czerw. 2025. Overview of respiratory sensor solutions to support patient diagnosis and monitoring. *Sensors* 25, 4 (Feb. 2025), 1078. doi:10.3390/s25041078

[11] Christos Kokkotis, Serafeim Moustakidis, Stefan James Swift, Flora Kontopidou, Ioannis Kavouras, Anastasios Doulamis, and Stamatios Giannoukos. 2025. Artificial intelligence and machine learning in the diagnosis and prognosis of diseases through breath analysis: A scoping review. *Information* 16, 11 (Nov. 2025), 968. doi:10.3390/info16110968

[12] Ik-Jae Kwon, Tae-Young Jung, Youjeong Son, Bongju Kim, Soung-Min Kim, and Jong-Ho Lee. 2022. Detection of volatile sulfur compounds (VSCs) in exhaled breath as a potential diagnostic method for oral squamous cell carcinoma. *BMC Oral Health* 22, 1 (July 2022). doi:10.1186/s12903-022-02301-3

[13] Massimiliano Laddomada and Massimiliano Pierobon. 2015. A crosstalk-based linear filter in biochemical signal transduction pathways for the internet of bio-things. In *40th International Conference on Acoustics, Speech and Signal Processing (ICASSP 2015)*. IEEE, South Brisbane, Australia, 5520–5524. doi:10.1109/icassp.2015.7179027

[14] Sajjad Mortazavi, Somayeh Makouei, Karim Abbasian, and Sebelan Danishvar. 2025. Exhaled breath analysis (EBA): A comprehensive review of non-invasive diagnostic techniques for disease detection. *Photonics* 12, 9 (Aug. 2025), 848. doi:10.3390/photonics12090848

[15] Stefan M. Moser, Po-Ning Chen, and Hsuan-Yin Lin. 2010. *Error Probability Analysis of Binary Asymmetric Channels*. Intermediate Report. Department of Electrical Engineering, National Chiao Tung University. https://moser-isi.ethz.ch/docs/papers/smos-2010-2.pdf

[16] Saswati Pal, Jorge Torres Gómez, Lisa Y. Debus, Regine Wendt, Florian-Lennert Adrian Lau, Cyrus Khandanpour, Malte Sieren, Stefan Fischer, and Falko Dressler. 2025. Machine Learning-Driven Localization of Infection Sources in the Human Cardiovascular System. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications* 11, 4 (Dec. 2025), 524–530. doi:10.1109/TMBMC.2025.3605770

[17] Anna Paleczek, Dominik Grochala, and Artur Rydosz. 2021. Artificial Breath Classification Using XGBoost Algorithm for Diabetes Detection. *Sensors* 21, 12 (June 2021), 1–18. doi:10.3390/s21124187

[18] George Preti, Erica Thaler, C. William Hanson, Michelle Troy, Jason Eades, and Alan Gelperin. 2009. Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: Analysis by solid-phase microextraction and gas chromatography–mass spectrometry. *Journal of Chromatography B* 877, 22 (July 2009), 2011–2018. doi:10.1016/j.jchromb.2009.05.028

[19] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. 2018. CatBoost: unbiased boosting with categorical features. In *32nd Conference on Neural Information Processing System (NeurIPS 2018)*, Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, and Nicolò Cesa-Bianchi (Eds.). Curran Associates Inc., Montréal, Canada, 6639–6649. doi:10.5555/3327757.3327770

[20] Christine Recinto, Theodore Efthemeou, P. Tony Boffelli, and James W. Navalta. 2017. Effects of nasal or oral breathing on anaerobic power output and metabolic responses. *International Journal of Exercise Science* 10, 4 (July 2017), 506–514. doi:10.70252/ehdr7442

[21] Adam J. Shapiro, Sharon D. Dell, Benjamin Gaston, Michael O'Connor, Nadzeya Marozkina, Michele Manion, Milan J. Hazucha, and Margaret W. Leigh. 2020. Nasal nitric oxide measurement in primary ciliary dyskinesia. A technical paper on standardized testing protocols. *Annals of the American Thoracic Society* 17, 2 (Feb. 2020), e1–e12. doi:10.1513/annalsats.201904-347ot

[22] Yasmim Carvalho Telson, Renata Maria Moreira Moraes Furlan, Matheus Pereira Porto, Rafael Augusto Magalhães Ferreira, and Andréa Rodrigues Motta. 2023. Evaluation of the breathing mode by infrared thermography. *Brazilian Journal of Otorhinolaryngology* 89, 6 (Nov. 2023), 101333. doi:10.1016/j.bjorl.2023.101333

[23] John B. West. 2012. *Respiratory Physiology: The Essentials*. Lippincott Williams & Wilkins.