ANALYSIS OF A MOBILE COMPUTING SYSTEM FOR INDOOR ENVIRONMENTAL MONITORING

by

JOSEPH CLEVENGER M◌LAUGHLIN

A THESIS

Presented to the Department of Computer and Information Science
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Master of Science

June 2021

THESIS APPROVAL PAGE

Student: Joseph Clevenger M⊆Laughlin

Title: Analysis of a Mobile Computing System For Indoor Environmental Monitoring

This thesis has been accepted and approved in partial fulfillment of the requirements for the Master of Science degree in the Department of Computer and Information Science by:

Michal Young	Chair
Siobhan Rockcastle	Core Member
Hank Childs	Core Member

and

Andy Karduna	Interim Vice Provost for Graduate Studies

Original approval signatures are on file with the Division of Graduate Studies of the University of Oregon.

Degree awarded June 2021

THESIS ABSTRACT

Joseph Clevenger McLaughlin

Master of Science

Department of Computer and Information Science

June 2021

Title: Analysis of a Mobile Computing System For Indoor Environmental Monitoring

Sensor networks that collect indoor environmental (IEQ) data are frequently used to drive building systems and to inform research and building standards. The research that connects the trends in IEQ data with human factors is well established and remains an active area of research. Conventional sensor systems often require specialized building infrastructure and are cost prohibitive, making them available primarily to large-scale indoor settings. In this thesis, we describe an alternative sensor network which utilizes a mobile sensor to collect IEQ data in traditionally under-served buildings. The sensor network is composed of inexpensive components and minimally relies on specialized infrastructure. We demonstrate that the sensor network maintains a uniquely fine spatial resolution, illustrating a unique utility beyond static sensor networks in conventional settings.

CURRICULUM VITAE

NAME OF AUTHOR: Joseph Clevenger M◌ͤLaughlin

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR, USA

DEGREES AWARDED:

Master of Science, Computer and Information Science, 2021, University of Oregon
Bachelor of Science, Computer and Information Science, 2017, University of Oregon

AREAS OF SPECIAL INTEREST:

Computational Science
High Performance Computing
Mobile Computing

PROFESSIONAL EXPERIENCE:

Senior Software Engineer, Kayhan Space, 2020-2021

Graduate Employee (Teaching) (Introduction to Web Programming, Fluency with Information Technology, Software Methodologies, Introduction to Python Programming), Department of Computer and Information Science, University of Oregon, 2020-2021

Software Engineer, Datalogic, 2017-2019

PUBLICATIONS:

McLaughlin, J., Young, M., & Rockcastle, S. (2021). Fine-grained measurement of the indoor built environment with robotic vacuum cleaners. *BS2021—17th International Conference of the International Building Performance Simulation Association.* Bruges, 13 September 2021 « forthcoming ».

Flores J., Fuentes R., **McLaughlin, J.**, Novitzky K., Schofield S., Springel A., Tal S. (2021). Pipeline Trees—An Auxiliary Tool in the Creation of Time Series Pipelines, *ITISE2021—International Conference on Time Series and Forecasting.* Gran Canaria, 19-21 July 2021 « forthcoming ».

ACKNOWLEDGEMENTS

I thank the members of my committee: Professor Michal Young, Professor Siobhan Rockcastle, and Professor Hank Childs for their assistance in the preparation of this manuscript. In particular I would like to thank everyone who so far participated in this research, including Siobhan Rockcastle and Michal Young, and Sam Peters who have demonstrated a commitment to cross-disciplinary research and have offered their continued support and assistance. Special acknowledgements are offered to Professor Anthony Hornof who assigned me a completed unrelated project that serendipitously kicked off this project.

For Mom and Dad and for my whole family who continue to give me their patience and encouragement.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

The study of *the indoor built environment*—with regard to its effects on those who occupy it—often involves simulation studies and controlled field experiments to argue for causality. Recently works have demonstrated that several attributes of the built-environment have measurable effects on occupants (Amundadottir, Rockcastle, Khanie, & Andersen, 2017; Bluyssen, 2013; Lan, Lian, & Pan, 2010). However the application of such research is limited to newly constructed buildings and primarily in "large-scale" settings where commercial environmental sensors and advanced environmental controls can be installed. Existing low-density homes and commercial settings where people spend a significant portion of their time indoors are unlikely to benefit from the application of this research without an analogous environmental sensor system. In this thesis, we discuss the development of a novel, low-cost mobile sensor. Through our analysis we demonstrate that our sensor system is both suitable for under-served locations and collects data has unique utility beyond the scope of traditional sensors.

1.1 Motivation

Conventional indoor environmental sensor systems comprise networks of environmental sensors (often measuring several environmental attributes), fixed within a building. Environmental sensors can either be wired directly into the building system or exist as part of an external system. Wireless sensor systems have also been developed and are often less expensive upon initial purchase; however, individual sensor packages have a fixed battery life and may be more expensive over time due to networking and maintenance requirements (Cree et al.,

2013). Sensor networks can drive environmental systems (e.g. heating, cooling, shading) or be used for diagnostic anomaly detection.

The motivation behind improving environmental sensing within the indoor built environment can be bifurcated into two concerns:

(i) **Energy Efficiency.** Heating and cooling comprises 55% of residential energy usage in the U.S. (U.S. Energy Information Administration, 2020). Residential energy usage continues to grow year after year as a portion of total U.S. energy consumption; in 2019 the sum of energy consumed by all U.S. residential buildings was approximately 1.2×10^{20} J, a figure that is 22% greater than the energy consumed by all commercial buildings (U.S. Energy Information Administration, 2020). Low-density residential buildings also have a greater *surface area* with more exterior walls to interior space compared to commercial and industrial buildings, further increasing the cost of heating and cooling. The U.S. Department of Energy (DOE) estimates that a series of simple and non-invasive infrastructure improvements to U.S. homes would reduce the cost of heating and cooling by 15% (U.S. Department of Energy, 2011).

(ii) **Occupant Health & Comfort.** People in upper middle-income economies spend a majority of their time inside human-built structures, some estimates placing it as high as 80%–90% (Baccarelli et al., 2011). Research has established several relationships between the health, comfort, and productivity of occupants and quantifiable elements of the indoor environment, including temperature (C), relative humidity (%), illuminance (lx), and particulate matter ($\mu g/m^3$) (Amundadottir et al., 2017; Bluyssen, 2013; Lan et al., 2010; Nezis, Biskos, Eleftheriadis, & Kalantzi, 2019).

Ideally, a building system should optimize for both of these goals, ensuring the comfort and health of occupants while meeting building performance goals. Environmental sensor systems can facilitate this balance by detecting faults within the environment or by influencing environmental controls to reduce waste. Sensor systems are often able to maintain the optimal temperature of a space where a human driven system might otherwise over-correct; however, building systems with this degree of reliability are often only found in commercial or industrial buildings. Small-scale residential and commercial settings often lack the necessary infrastructure.

1.2 Constraints

Conventional environmental monitoring systems are the product of industrial motivations, designed primarily for corporate and industrial settings where minor shifts in building performance can represent a large and easily identifiable cost. Some tools described as *research-oriented* are available for purchase, but these tools are often expensive and inaccessible. In both cases, the range of available sensing systems broadly under-serves many environments; where a sensing system can be deployed, sensors must be fixed onto the environment, requiring assumptions about occupant behavior and performance patterns. The development of our mobile sensor is primarily motivated by two limitations of conventional sensor networks:

(i) **Performance constraints.** As they are fixed to physical locations, conventional sensor networks have a spatial resolution that is proportional to the number of sensors in the network. The capacity of the sensor network to detect faults or guide building performance depends upon *where* the sensors are fixed. Recent work has demonstrated how appropriately positioning

sensors can significantly improve fault detection tasks (Fernandez et al., 2017). However, even a system performing within the parameters of acceptable building performance requires several assumptions about occupant behaviors and furniture layout. This suggests that there are a complex orchestration of assumptions required to compose fixed sensor readouts to dynamic indoor spaces.

(ii) **Deployment constraints.** Conventional sensor networks are most common in large commercial and industrial buildings; smaller residential and commercial buildings rarely have analogous sensor networks. The absence of complex sensor systems in residential buildings can be attributed to the individualized costs associated with these changes. Leaving this class of buildings under-served ignores the potential effects of these environments on their occupants as well any reductions in energy usage.

1.3 Thesis Statement

In this thesis, we perform an analysis of a mobile computing system designed to measure the built-environment in a uniquely granular manner. Our system collects indoor environmental quality (IEQ) data on a roaming mobile device. This data has a considerably variable spatial resolution, finer than conventional sensor systems while at the same time only requiring a single sensor package. Our analysis captures both the construction of the device as well as a pilot study designed to establish a baseline of the device's performance. Through this analysis we show that the sensor system demonstrates improved spatial resolution over existing sensors and is well-suited to traditionally under-served locations.

1.4 Contents

Our analysis is preceded by chapter II (Background) in which an overview of the mobile sensing system is introduced and our design objectives for the system are discussed. Chapter III (Methodologies) details the procedures we followed to construct and evaluate our mobile sensor, the construction of the separate components in the completed system (e.g. physical components, software components, computational methods), as well as the design of a pilot study from which we collect an initial set of data. Chapter IV (Analysis) provides our analysis of the data collected from the study. We define and assess both *the performance and viability* of the device as well as *the direct inferences* that can be made from the results. Following our analysis, chapter V (Related Works) compares our device and our computational methods to other contemporary works in the area. We conclude with a review of our analysis in chapter VI (Discussion), where we evaluate the limitations of our approach and its implications for future work.

CHAPTER II

BACKGROUND

2.1 Overview

Our sensor system is distinguished from conventional sensor networks in one primary way: instead of depending upon a collection of fixed sensors, we depend upon a single mobile sensor package. Here, a mobile sensor is one that has some degree of mobile autonomy within the space. With this mobility we are able to collect measurements across a highly variable spatial domain. Where conventional fixed sensors would have to make assumptions about the data missing at the locations between their fixed positions, a mobile sensor can directly observe those locations. This distinction is illustrated in Figure 1.

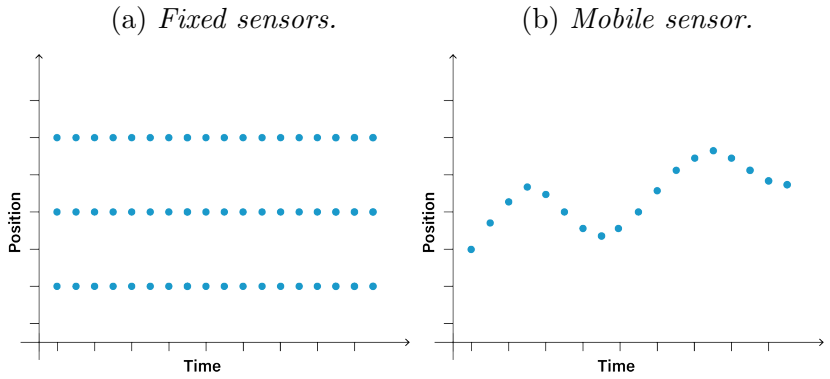
From a high level, we accomplish this with a collection of inexpensive and commercially available hardware components. The device is mounted on top of an autonomous robotic vacuum cleaner. We attach a series of environmental sensors to the vacuum cleaner and largely allow the vacuum to operate normally. To determine the location of the sensor we use an external location estimation network. A small, single-board computer coordinates the system.

2.2 Objectives

Our mobile sensor addresses the limitations we identified in existing, conventional fixed sensor systems previously identified in §1.2 (performance constraints and deployment constraints.) To address these constraints, we establish three design objectives:

(i) **Improved spatial resolution.** The sensor system should improve upon the spatial granularity of fixed sensor networks.

Figure 1. Fixed sensor and mobile sensor data illustrated.



An illustration of the differences between data collected by fixed networks (a) and data collected by mobile networks (b.) For the purposes of this illustration, "position" has been reduced to a single axis while in practice it may comprise 2-D or 3-D coordinates.

(ii) **Minimized interference.** The deployed system should minimally disrupt the existing environment. Further, it should seek to utilize common elements within the indoor built environment.

(iii) **Flexible deployment.** The sensor system should be suited to several indoor settings, in particular those under-served by conventional sensor systems.

The core of our design—the implementation of the vacuum as a *host* for the sensor system—addresses the first two objectives. The vacuum is inherently mobile, collecting data with a fine spatial resolution; the trade-off of a single mobile sensor is the loss of *consistent* spatial data. The choice to use an autonomous vacuum as the host for the sensor system is an explicit choice that emphasizes the utilization of objects common to under-served residential buildings. This choice establishes a dependence upon existing systems instead of developing a custom system that may be more optimal. In the subsequent chapters we evaluate these trade-offs and how our implementation fulfills these objectives.

CHAPTER III

METHODOLOGIES

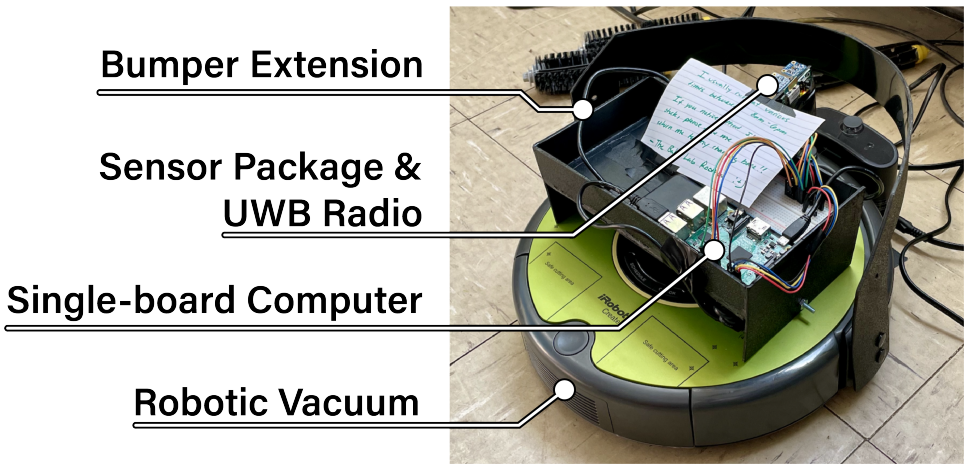
This chapter details the methodologies we developed in the construction and analysis of our mobile sensor system. We begin by introducing the *design goals* of the device, and how the state of the application area motivates these goals. Then, we describe the device's construction (both physical and software) as well as a pilot study designed to evaluate the operational capacity and utility of the system. Finally, we detail the computational methods we derived to evaluate the data collected in the pilot study.

3.1 Physical Construction

We selected a collection of common hardware components to compose the final device. At a high a level, the construction of the device consists of a common single-board computer, integrated with several inexpensive commodity sensors, all mounted on top of a robotic vacuum cleaner. A photograph of the assembled device is shown in Figure 2.

The component that illustrates our design objectives is the vacuum cleaner. We selected this specific device (iRobot Roomba 600) because it is representative of a class of devices (i.e. robotic cleaning devices) that are common in residential, as well as commercial and industrial settings. In a broad sense, the vacuum acts as a *host* for the sensor system, demonstrating that we can deploy such a system without dramatically altering the configuration of the indoor environment. This specific device is also robust, inexpensive, and the particular version we used came with a programmable serial interface, which we utilized for remote control of the vacuum's operations.

Figure 2. Labelled photograph of the device.



An image of the physical device we constructed, including the single-board computer, sensors, ultra wide-band radio, robotic vacuum cleaner, housing, and bumper modifications.

3.1.1 Evaluation of Materials. Prior to the assembly of the device, we performed an assessment of the different components we selected. The vacuum was of particular interest, being a commercial product and fully functional in its own. Our assessment considered the capabilities and limitations of the device that may influence our assembly procedure.

The details of the vacuum's pathing algorithm are proprietary; however our assessment of the vacuum suggests that the algorithm largely just reacts to obstructions in its environment, and upon an obstruction it selects a new heading. The device primarily detects environmental obstructions through an array of both optical sensors and physically actuating sensors, mounted on the *bumper* section of the vacuum.

We evaluated the vacuum's odometry through the vacuum's serial interface where we initiated a series of `move` and `turn` instructions. The specific instructions

can be found in the manufacture's specification.[1] For our purposes, we modified the speed of the motors' running the vacuum's wheels proportionally, so that we either measured the distance travelled in a straight line, or a turn performed with a *zero turning radius* (i.e. rotating about its central axis.) In general we found that the measurements taken *ex post* from the `move` varied on the order of 20 cm and measurements taken after the `turn` instruction varied on the order of 3° from the intended action. Error on this scale made the vacuum's odometry information unsuitable for our goal of reliably estimating the device's location.

3.1.2 Ultra Wide-Band Networks. Ultra wide-band (UWB) radio devices have emerged as a mechanism for short-range communication (i.e. within the range of 1—100 m.) *Ultra wide-band* broadly applies to radio spectrum communication with a bandwidth >500 MHz; the UWB designation originates with the U.S. Federal Communications Commission (FCC) though is commonly used to refer to high bandwidth radio devices for short-range communication.

While this remains an active area of active research, UWB networks for location estimation are broadly based on the principle of Time-of-Arrival (ToA) estimation (Falsi, Dardari, Mucchi, & Win, 2006). Under a ToA estimation regime, location is derived from the time taken for a message to reach its destination through a medium (e.g. radio waves) from a known source location. Using at least 3 known sources, the position of the destination can be derived. Other factors may impede the location estimation task, such as the speed of the medium; though effectively constant through air, the speed of UWB communication may vary through other mediums. UWB is particularly suited to this task because the bandwidth is great enough to encode the information necessary for a ToA regime

[1] `https://www.irobotweb.com/~/media/MainSite/PDFs/About/STEM/Create/iRobot_Roomba_600_Open_Interface_Spec.pdf`

(i.e. source location, emission timestamp.) Research into UWB location estimation through similar methods has made cm-level indoor positioning possible (Zafari, Gkelias, & Leung, 2019).

To determine the location of the host device, we used an existing UWB-based location estimation network (Decawave, 2020). The network is composed of at least three UWB transmitters configured as *anchors* (i.e. radios with known locations) and at least one UWB transmitter configured as a *tag* (i.e. a destination where the ToA is calculated.) The devices are configured and polled through a serial interface. In practice we found that the network accurately determined the location of the tag within a range of 10–15 cm.

3.1.3 Miscellaneous Components & Assembly. We constructed the physical system using a single-board computer, inexpensive commodity sensors, and a location estimation network composed of UWB radios. The computer, sensors and UWB radio are all attached to the top of the vacuum. The vacuum contains programming to allow users to schedule it to navigate and clean the room once a day, but we utilized the vacuum's exposed serial port to remotely initiate the cleaning cycle from the single-board computer. Though our goal is to minimally disrupt the typical configuration of the built environment to enable our device, we must accept that this arrangement adds an artificial presence to the space. The device has the capability to run the vacuum's cleaning cycle for longer durations and at greater frequencies than a user would expect; however, we accepted these limitations as a part of this design.

For our implementation we used a Raspberry Pi single-board computer and DWM1001 UWB radios (Decawave, 2020). We attached the anchor radios to the walls of the space and the tag radio to the vacuum. The tag radio interfaces with

the computer over a USB (virtualized serial) interface. UWB radios operate best with line-of-sight communication, so we elevate the anchors 1 m above the floor to increase the likelihood that line-of-sight is maintained between the anchor and tag. We attached two sensors to the computer, a `BME280` which captures the illuminance (lx) and a `TSL2561` which captures air temperature (C) and relative humidity (%.) Both devices communicate with the computer over a general purpose input-output (GPIO) interface.

The physical structure of the vacuum was modified to accommodate the components attached on top. An elevated tray was attached to the top of the vacuum to store the computer and a lithium-ion battery which acts as the computer's primary power source. We raised the tray to not interfere with the manual button interface of the vacuum located on the top of the vacuum. An antenna arm extends off of the tray to anchor the environmental sensors and the UWB tag.

Since we are utilizing the vacuum's pathing algorithm, we had to modify how the vacuum registers collisions to account for the components attached on top. To do this, we built and attached a vertical extension to the vacuum's bumper. The extension is composed of heat moulded ABS plastic that allows for collisions with the extension to be communicated as collisions to the vacuum. Without the extension, the components mounted onto the top of the vacuum would likely collide with objects in the environment without communicating the collision with the vacuum. Figure 2 shows an annotated photograph of the final device.

3.2 Software Systems

We developed a suite of software responsible for the concerns of: (i) the host, (ii) remote data storage and retrieval, and (iii) data processing, analysis,

and visualization. We largely adhered to common industry tools, developing the majority of the software in `Python3`. This section details the capabilities and constraints each of these three systems.

3.2.1 Host Software. The host software system is deployed locally on the single-board computer and coordinates all connected components (GPIO sensors, ultra wide-band radio, vacuum), as well as with the remote data storage system. Each component reads-out information at a different rate, and with a different reliability; necessarily, the host software system gives special consideration to the constraints of each individual component. Generally we found that the ultra wide-band radio had the smallest latency (on the order of milliseconds;) the ultra wide-band radio also reports a *confidence interval* that rates the confidence of the estimated location. We saw that if the host navigated too far beyond of the line-of-sight of the anchor radios, the confidence interval dropped significantly. Without line-of-sight, high confidence location estimations still occurred, but were less frequent. The latency of the GPIO sensor readouts for illuminance, humidity, and temperature were unequal, though all on the order of $1{\times}10^3$ ms.

To accommodate the variability between GPIO sensors and the UWB radio, we poll all GPIO sensors at once and any attempted sensor readout that captures at least one sensor readout is paired with the next location estimation with a confidence $\geq$50%. Paired readouts are added to a thread-safe queue, to be consumed by a separate thread that dispatches the results to the remote data storage system. Calls to the remote system can vary in duration, dependent on the state of the network; the interactions with the remote system are placed on their own thread to prevent the main thread from blocking on slow network calls. This process occurs on a loop, dependent upon some time threshold, δ s.t. $\delta > 0$.

The loop condition executes until the current time t, exceeds the halting time, $t_0 + \delta$ where t_0 is the initial time. The whole process thus far is preceded by a `clean` command from the single-board computer, which initiates the vacuum's cleaning cycle. Similarly, a `seek_dock` command is sent to the vacuum following the completion of the cleaning cycle; the `seek_dock` command is sent twice, once to terminate the cleaning cycle, and again to initiate a procedure in the vacuum's programming to seek out its charging dock. This process is illustrated in Algorithm 1. This system was written in `Python3` and was dispatched by configuring the single-board computer's `crontab`.

3.2.2 Remote System Software. We designed the remote data storage system with a few requirements. Primarily, that the system should maintain asynchronous access (for both creation and retrieval) to the data created by the host software. Further, the system should support several host devices in a manner that can be viewed as *generic*. Generic in this sense refers to the capacity of the system to accept input indiscriminately. These requirements are fairly common for modern web services, as is the methodology we follow is development of this system.

The system comprises an web representational state transfer (REST) API, accessible to both the host software and external data consumers. The API integrates with a database for persistent storage (in practice we chose `postgres`.) The API was written primarily in `Python3` using the `Flask` web framework library. We deployed the system on Google Cloud Platform, which manages the *scale* of the application in the event that several hosts or consumers make API calls to the system at once.

We categorize every spatiotemporal measurement taken by the host system as an *annotation*. Every annotation is associated with a *collection* that represent an instance of the host system collecting new measurements. The database field, `annotation` has a column for the associated collection the location and time of the measurement taken, as well as a `blob` data column for the associated measurements. The specific sensor readouts are variable, and leaving this field flexible suits both the variability of the current implementation and supports additional sensor data in future devices.

Algorithm 1 Host software system.

Require: Initialize *queue, δ, running*

procedure MAIN THREAD
    token ← *login(remote)*
    if *token* = NULL **then return**
    write(SerialPort, CleanCommand)
    *launch(*NETWORK THREAD*)*
    epoch ← *systime()*
    halt ← *epoch* + *δ*
    while *epoch* < *halt* **do**
        readouts ← *read(GPIO)*
        if *readouts* ≠ *NULL* **then**
            position, confidence ← *read(SerialPort)*
            while *confidence* < 0.5 **do**
                position, confidence ← *read(SerialPort)*
            push(queue, pair(epoch,position,readouts))
        epoch ← *systime()*
    write(SerialPort, SeekDockCommand)
    running ← false
    *wait(*NETWORK THREAD*)*
    return
procedure NETWORK THREAD
    while *length(queue)* > *0* or *running* **do**
        pair ← *pop(queue)*
        POST(token, pair)
    return

Access to the system is authorized with a JSON web token. Upon a successful login, the device (host or consumer) is passed a token from the remote service that is exchanged in the HTTP header of messages from the device to validate its origin. Though there are other mechanisms for authorizing REST calls, JSON web tokens are fairly simple to implement. A downside to JSON web tokens is that the token must be exchanged with every REST call, adding a bandwidth overhead with every message.

3.2.3 Analysis Software. To evaluate the data created by the host software we developed a collection of analysis tools. These tools are primarily developed in `Python3` and facilitate the requisition of data stored on the remote system, and the analysis and visualization of that data.

Data from the remote system is collated by its collection and saved in a static data format. Analysis tasks may involve several collections, as collections may only span minutes, so the analysis software was developed with the capability to consume multiple collections for a single analysis. We include several libraries to support different inference and prediction tasks including `numpy`, `scipy`, and `scikit-learn` (Harris et al., 2020; Pedregosa et al., 2011; Virtanen et al., 2020). For visualization tasks we included two mechanisms: a real-time visualisation scheme using `matplotlib`, and an open vector graphics format for high-resolution, post hoc visualizations Hunter (2007).

3.3 Pilot Study

To evaluate the performance of the device we designed a pilot study. Performance captures the aspects of the system's functionality and reliability, as well as the quality of the data that we are able to capture. We evaluated whether

or not the device experienced any operational failures as well as whether the data collected in the study is suitable for prediction or inference.

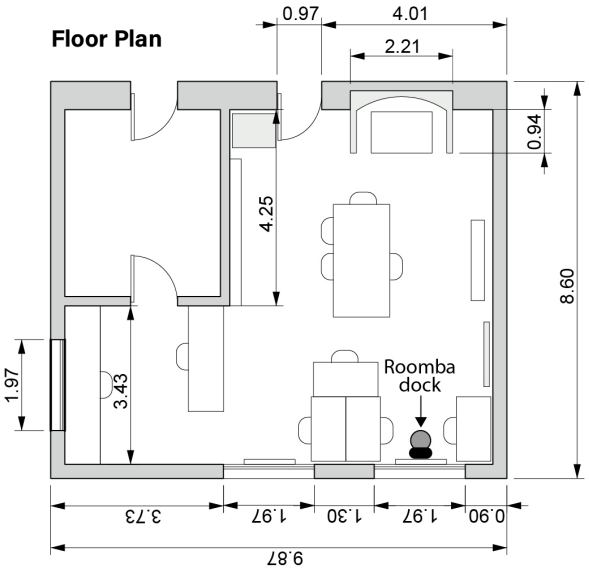
Given our focus thus far on the practical nature of the device, our choice of environment was one that represented a more natural approximation of a real indoor environment. We piloted the device in a laboratory space, approximately 70 m^2, at The University of Oregon. The room was on the second floor of the building and was fitted with two south facing windows. The room also had a west facing window, but this window was obscured for the duration of the study. Below the two south facing windows were a pair of wall-mounted heaters. The room was also equip with ceiling mounted air circulators. The space was laid out in with several desks, chairs, and other laboratory equipment; the furniture and equipment, as well as the frequent occupants of the space, provided a suitable representation of a "real-world" indoor space. A floor plan of the space is illustrated in Figure 3.

The device's charging dock was fixed at the south-east corner of the room, against the southern wall. Three UWB radios were configured as *anchors* and fixed to the walls of the room, approximately 1 m above the floor. The network was configured such that the south-east corner of the room was fixed as the 3–axis origin (i.e. $(0, 0, 0)$) and all anchors were configured with their location relative to the origin.

3.4 Computational Methods

Time-series data using parsed, fixed sensors is well studied in fields such as Geographical Information Systems (GIS) where it has been utilized from a host of applications such as habit monitoring and flood monitoring (Castillo-Effer, Quintela, Moreno, Jordan, & Westhoff, 2004; Mainwaring, Culler, Polastre, Szewczyk, & Anderson, 2002). The data collected by the mobile sensor is unique

Figure 3. The floor plan of the room used in the pilot study.



The floor plan of the approximately 70 m^2 *space used in the pilot study. Location of the vacuum's docking station is labelled; furniture layout is illustrated; measurements are given in meters.*

from other common sensor time-series data in that it is collected by a mobile device; this means that there will only be value at any point in time, but over longer time intervals the data over space can emerge to cover the entire area at a high resolution. This data is distinct from data collected in fixed sensor networks and requires a distinct method of evaluation.

This has been viewed previously as an optimization problem wherein the goal is to find a function that best estimates the value of an unexplored location at a certain time, given a time series of locations and values

$$\mathcal{T} = \{t_1, p_1, v_1\} \dots \{t_n, p_n, v_n\}$$

where t_k, p_k, v_k are time, location, and value accordingly (Jin, Liu, Schiavon, & Spanos, 2018). The authors' use a common prediction model (e.g. random forest, K-NN) to extract *local trends* in the spatiotemporal data, then refine

this prediction with the residuals of a linear model. The details of this method are discussed in §5.1. This assumes that obvious global trends emerge in spatiotemporal IEQ data. An example of this intuition is such: *on cold days, a room is likely to trend cooler than on warm days.*

Our intuition challenges this notion with the following: *on cold days, the presence of the heater will manifest as a distinct "lumpiness" of temperature within the room.* Anecdotally, we experienced this in space used in the pilot study; the lower-left corner of the room was consistently cool and the locations near the window were consistently brighter.

Another distinction between the authors' conditions and our own, is that our mobile sensor travels on a path that is completely opaque to us. Jin et al. (2018), however used a device with fairly deterministic path circling around a room.

We model the *regional trends* we observed, deriving residuals from a K-MEANS classifier, training on every attribute as a predictor (James, Witten, Hastie, & Tibshirani, 2013). The residuals refine a predictive model using using the measured quantity as a response. The procedure for this method follows:

(a) Train a K-MEANS classifier using the spatial, temporal, and measured values as predictors,

$$\hat{Y} = \text{KMEANS}(\mathcal{T})$$

where $\hat{Y}$ is a classification with respect to one of j. We use j to refer to the k clusters we select in K-MEANS to not conflict with subsequent usage of k in K-NN. In practice, we chose to optimize j with the silhouette method (Rousseeuw, 1987).

(b) All spatiotemporal values sharing the same classification $\hat{y}_i \in \hat{Y}$ are fitted with LOWESS regression,

$$f_i = \text{LOWESS}(y_i).$$

The residuals $r_i, i \leq j$ are determined from the LOWESS fit line for that class

$$r_i = f_i(y_i) - y_i.$$

(c) The value at a given spatiotemporal location t, p, is determined by training weighted K-NN with spatial, temporal, and values as predictors and measured values as the response.

$$\hat{\hat{y}} = \text{KNN}(t, p).$$

Post hoc, $\hat{\hat{y}}$ is refined with the residuals from LOWESS smoothing.

This allows the residuals from multiple regions to influence a particular inference, particularly when samples come from several classes and if k is large.

CHAPTER IV

ANALYSIS

In this chapter we evaluate the result of our pilot study. The parameters we fixed in the study discussed, as well as the performance of the space and systems we utilized. Then, we consider the results we collected, explore any trends that emerge, and the validity of those trends. Through our analysis we demonstrate fulfillment of our design objectives defined in §2.2.

4.1 Pilot Study

The pilot study occurred over five days, from 6 January to 10 January 2021. Summing all 20 minute collection periods, the device collected 10 hours worth of data. The periods occurred on 2 hour intervals, beginning at 8:00 and ending at 16:00 everyday. We recorded approximately 70 000 illuminance measurements and 109 000 temperature and relative humidity measurements. The discrepancy between the total number of measurements is due to hardware limitations (i.e. readout rate) in the different sensors. Illuminance measurements are consistently read at a lower rate than temperature and relative humidity. Each 20 minute collection period consists of 2 000–4 000 annotated measurements, resulting in a dataset of the measured attributes across space and over the time interval of the collection period.

4.1.1 System Performance. Each consecutive day saw similar performance, with approximately 20 000–26 000 measurements taken per day. The notable exception was the first day, 6 January, where the device encountered obstructions during two its six cleaning cycles; the cause of the obstructions were determined to be a low battery in the vacuum, as it did not return to its dock at the end of the previous collection period. Table 1 displays the daily totals over the

Table 1. Total daily measurements recorded during the study

Date	6 Jan.	7 Jan.	8 Jan.	9 Jan.	10 Jan.
Total Measurements	11 918	20 214	26 269	23 916	26 160

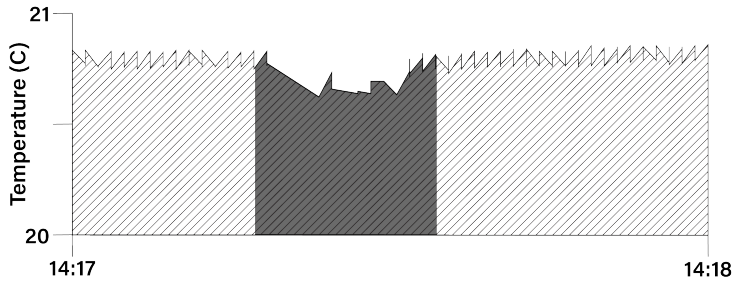
The number of measurements recorded by the device during the pilot study, totaled by the date; typically 20 000–26 000 measurements are taken per day except for the first day, 6 Jan. 2021, which recorded significantly fewer; the variance in the number of measurements taken on 6 Jan. 2021 was due to the device encountering obstructions during two of its cleaning cycles.

entire study. For the purposes of this analysis, we include the obstructed collections in the 6 January result when evaluating trends over the data. While the total daily measurements for 6 January is low, it is still below 2 (1.82) standard deviations (5 358) from the mean (21 695.)

The variance in the remaining daily totals can be attributed to uncertain location estimations that are rejected by the host system. The UWB radio's mechanism for determining the confidence of a prediction employs an internal uncertainty mechanism which often benefits from momentary acceleration to determine a more certain location (Li et al., 2017). In some instances, the host happens to travel in steady, straight paths reducing the number of high confidence location estimations; instances where the host collided with several objects in its path or performed turns is likely to cause more high confidence location estimations. Given this assumption, the varied frequency of measurements in the results should not significantly influence our analysis because they represent a greater frequency of measures in recently visited regions.

Further variance can be attributed to a lack of line-of-sight in some areas of the room. *Discontinuities* emerge when the UWB radio is completely

Figure 4. An example of a data discontinuity, plotted as a time-series.



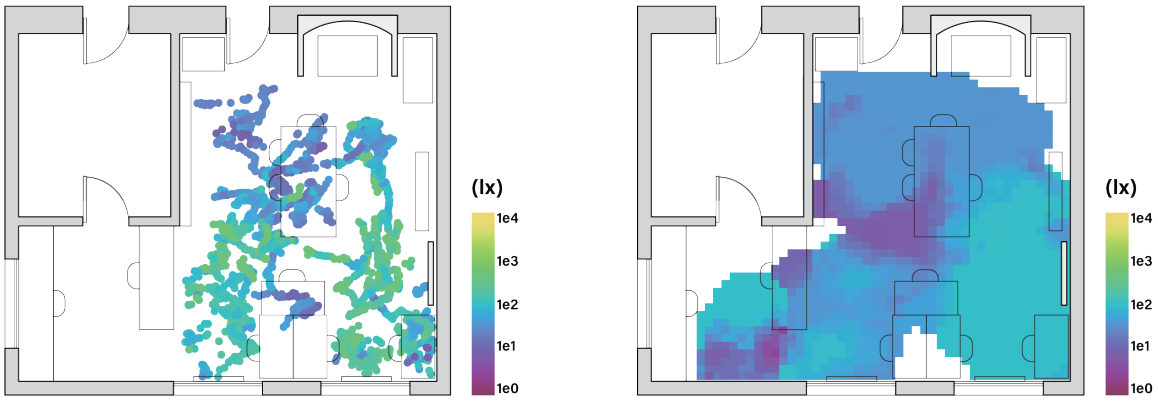
Temperature plotted over time from 14:17 — 12:18 local time, 8 January 2021; this plot contains a data discontinuity caused by low-confidence location estimations.

obstructed in its line-of-sight from the remaining anchors in the network; line-of-sight obstructions most commonly occur when the host navigates below furniture. Figure 4 plots a data discontinuity in temperature data caused by the device moving below a desk. This discontinuity appears as a gap in the frequency of points in the center of the plot. Notably, the location estimation network is still able to determine the position of host without line-of-sight, but at a slower rate. Also notable is that the measurements taken during the discontinuity are very slightly cooler than the neighboring measurements, suggesting that the air temperature under the desk was cooler.

Figure 5a illustrates the raw, annotated measurements over a single collection period. Every circle on the plot represents a single measurement in space. Every measurement is captured at a unique point in time within the 20 minute window, though time is not communicated in this figure. In this instance the vacuum did not reach the *nook* in the lower left corner of the room in its 20 minute period. We found that this area of the room was missed in roughly half of all collections. We attribute this to complexity of the room, and in particular because this area is only accessible by an approximately 1.5 m wide path. Further,

Figure 5. Raw and interpolated illuminance measurements plotted.

(a) *Raw.* (b) *Interpolated.*



Illuminance as measured over space from a single collection period (12:00 — 12:20 local time, 7 January 2021;) raw measurements (a) are plotted at their annotated location; the result of interpolating with our regional interpolation method to 12:00 (b;) log scaling used for lux.

the device's dock is located on the other side of the room, making it the most remote location in the space.

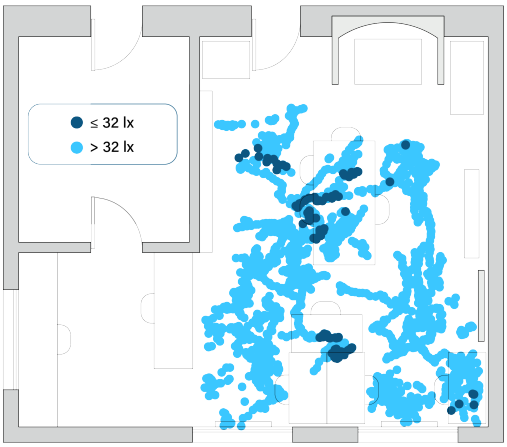
4.1.2 Data Properties. The effects of daylight at noon can be seen through the raw measurements in Figure 5a. Bright illuminance measurements are taken near the windows of the room and become less bright as distance from the window increases. This relationship is particular to a north/south-facing window; east/west-facing windows will typically allow more light at morning/evening hours.

On occasion, illuminance measurements display *artifacted features*—where significant variance occurs for short intervals during the collection period. This can be seen in Figure 6b. In general, brief and extreme variance in illuminance measurements can occur for two reasons: (i) the natural variability of daylight and (ii) the dynamic interactions between objects in the environment. Greater values that appear varied (i.e. in the range of 1×10^3—1×10^4 lx) may be more likely an effect of the variability of daylight. The effects of daylight are more notable

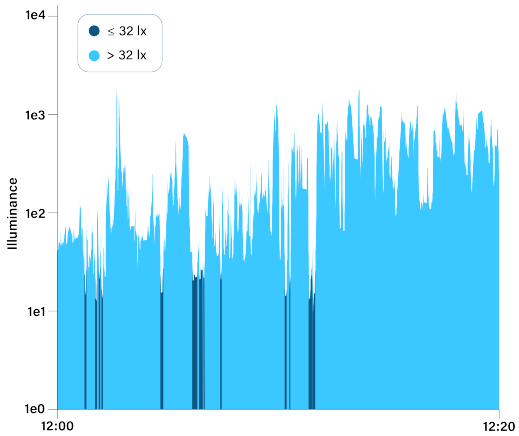
in Figures 5 & 6 because they were captured at 12:00. However, extremely low and varied values (i.e. 0–50 lx) may be more likely an effect of host's interactions with the environment; instances of the device traveling through the shadow of an occupant or under furniture would account for exceptionally low illuminance values. We demonstrate this relationship in Figure 6, which highlights the locations where illuminance measurements of less than 31.62 (i.e. $1{\times}10^{1.5}$) lx were recorded. The locations where illuminance is highlighted in Figure 6a are inside of the known furniture objects labelled on the floor plan.

Figure 6. Fixed sensor and mobile sensor data illustrated.

(a) *Floor plan plot.*



(b) *Time-series.*



Illuminance as measured over space from a single collection period (12:00 — 12:20 local time, 7 January 2021). The plots are bifurcated by values greater than 31.62 *(i.e.* $1{\times}10^{1.5}$*) and less than* 31.62*. The latter consists of locations that are below desks, chairs, and tables.*

Evidence of the relationship between the data discontinuities, artifacted features, and obstructions with furniture is shown in Figure 7; Figure 7 emphasizes an artifact that emerges in the both illuminance and temperature data. The location of these measurements reveal that they occurred in the desk cluster in the lower portion of the floor plan. Here, it is also clear how the artifact emerges

Table 2. Mean hourly illuminance on 7 January.

Time of Day	8:00	10:00	12:00	14:00	16:00	18:00
Mean (lx)	14.8	49.7	284.3	302.3	61.1	10.0

differently depending on the sensor readout: Illuminance drops to near zero almost immediately while temperature trends down in a continuous manner.

Discussion of the variability and properties of the dataset, as well as additional auxiliary figures is continued in Appendix A.

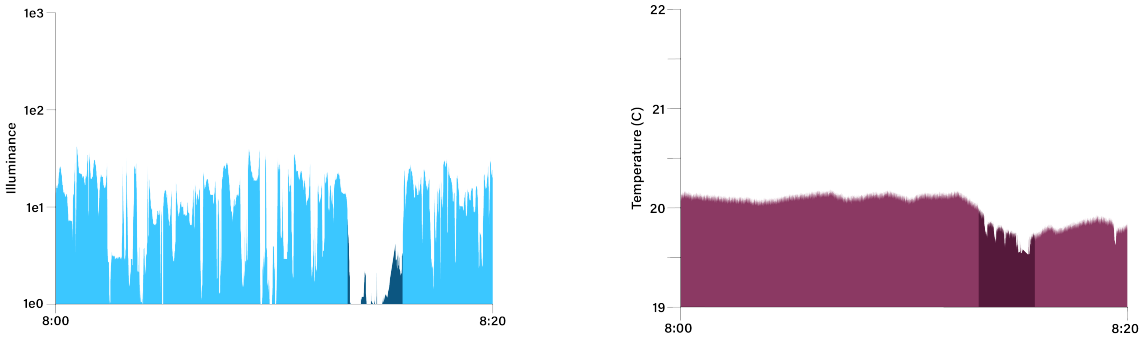
4.2 Interpolation Performance

Figure 5b shows the result of interpolating with the K-MEANS residuals smoothing method; the method is fixed to a 10 spatial discretization and is solved for 12:00, 7 January. Most notable here is that the result recovers values in the nook, where none were taken in the raw readout. We accomplish this by interpolating over space and time, and derive predictions for this area from measurements taken in neighboring collection periods. This allows us to recover lost measurements, but may produce a result that is less reliable. Between 12:00—14:00 often represented a *peak average* in bi-hourly illuminance measurements, this is shown in Table 2. As the averages of the prior are hours are likely to be lesser than those taken at noon and subsequent hours may be similar, the results that we recover may not be highly representative.

In Figure 5b, we have limited our predicted values to those with a Euclidean distance of less than 30 cm from a sampled location; though we could feasibly predict *any* spatiotemporal position, predicting positions that are sufficiently distant from the K-MEANS produces striated values with a limited predictive capacity. Further, the only valuable positions are those within the bounds of the

Figure 7. Time-series of illuminance and of temperature.

(a) *Illuminance* (b) *Temperature.*



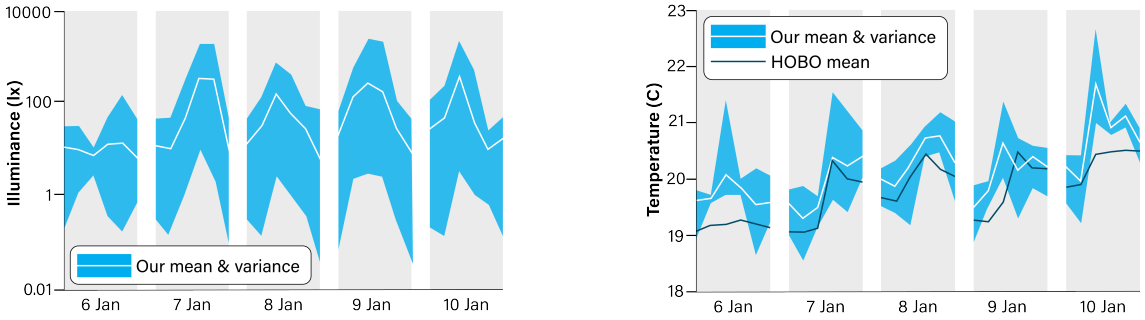
Illuminance (a) and temperature (b), plotted as a time-series as measured over space from 8:00—8:20 local time, 8 January 2021. The analogous section artifacted features are emphasized in each plot.

physical space. Limiting predictions to those that are near sampled locations does not guarantee form in our predictions (i.e. that the predicted locations fall within the bounds of the physical space,) but it does reduce the likelihood of predicting values with highly uncertain data.

4.2.1 Predictive Experiment. To evaluate the performance of the regional interpolation method we evaluate the illuminance and temperature data collected on 8 January. We split the data into a training sets and test sets for each respective response and used the following predictors: x-axis position, y-axis position, date-time. The K-MEANS model and the K-NN were trained only on the test set and evaluated and the result was evaluated on the test set. The only preprocessing on the data was to convert the illuminance measurements to `log10` space. This was done to allow sufficiently large illuminance measurements to contribute to the model. These values were incidental enough that randomly sampling the data did not result in representative splits. When moved to log-space, the splits were more representative..

Figure 8. Hourly illuminance and temperature, plotted against HOBO mean.

(a) *Illuminance.* (b) *Temperature.*



Illuminance (a) and temperature (b) (mean and variance) taken over 5 days (8:00 — 18:00) from 6 January 6 to 10 January 2021. Summed over each collection period (log scaling used for lux.) Includes HOBO fixed sensor readout.

The models predicted illuminance with a root mean square error (RMSE) of 0.93. Illuminance was converted to log-space space so the error term must be understood in orders of magnitude in lx. The temperature model performed with a RMSE of 1.06 on the test set.

4.2.2 Comparative Analysis. Concurrent to the study, we fixed six HOBO data loggers in the space. These devices are small battery powered sensors capable of measuring IEQ data for weeks at a time. Their application yields an approximation of a fixed sensor network to which we compare our mobile sensor data with. The fixed sensors were located on walls, chair legs, and desks across the room, and approximately 15 cm above the floor to put them at the same height as the mobile sensor system. Their orientation did not gather illuminance data with a similar profile to the mobile sensor's illuminance data. *Illuminance* is the measure of incidental light that strikes a surface, so this result is not surprising. As the mobile sensor travels throughout the space, it was able to collect a robust

profile of incidental light as it behaves dynamically in throughout the space,, not only on walls and surfaces of the room.

In Figure 8 we plot the daily averages of illuminance and temperature over the five days of the study. Here, the most notable anomaly is in Figure 8a where the first day has a distinctly different shape than the rest of the days. While 6 January was the day where two obstructions in the collection process occurred, these obstructions do not account for this pattern. We note that on 6 January, the weather was particularly cloudy and may have contributed to the difference. Figure 8b plots a similar series of temperature means, alongside the fixed sensor means. This figure is notable because the mobile sensor's means are consistently greater than the fixed sensor means. This is attributed to the mobiles sensor's fine spatial resolution that allows it to directly record the values at various heat sources. A fixed sensor system would require assumptions about the space, such as the location of a heat source, to make predictions about its behavior.

4.3 Distribution & Reliability

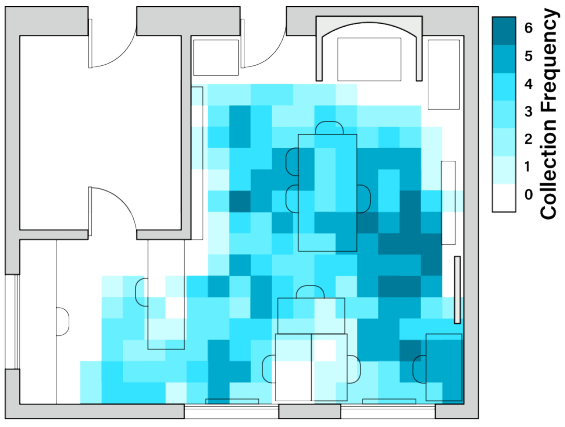
Utilizing the vacuum as our navigational host to the mobile sensor has the inherent quality of introducing uncertainties in the results. The host system does not control where the host travels to, only when the host should begin and cease travelling. Often, the 20 minute collection periods did not provide enough time for the vacuum to circumnavigate the entire space, as demonstrated in Figure 5 where the lower-left nook was not visited. Our observations conform with the assumption that spaces with fewer obstructions and more open spaces are more likely to be represented in the results. Figure 9 illustrates the frequency of measurements taken with the mobile sensor, with respect to location (discretized into a 50 cm^2 grid) across different collection periods and within a single day. Frequency in this context

increases when measurements are taken at the same location, but across different collection periods, excluding repeated sampling within the same period. In this figure we only count a grid square as "visited" if at least ten measurements were taken within its bounding box.

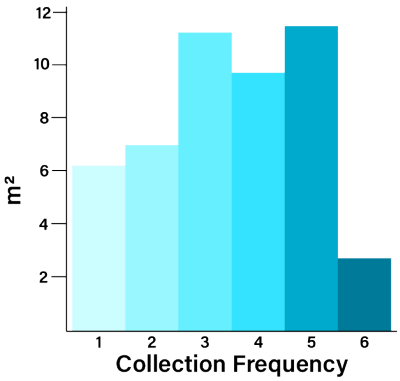
From Figure 9 we see that 49% of the *navigable regions* (those with floor access) appeared in at least three of six collection periods and that 29% of all navigable regions appeared in at least four of six collection periods; further, the regions that appeared in at least five collection periods are thoroughly distributed about the room, suggesting the host *is* traversing the room with some degree of a normal distribution within any given collection period. The exception to this is the lower-left nook, which is both narrowly obstructed and the furthest path from the host's docking station.

Figure 9. Frequency of measurements taken on a single day, plotted.

(a) *Overlay plot.*



(b) *Histogram.*



The frequency of measurements taken from 50 cm^2 regions of the space, over a single day (6 total collection periods on 8 January 2021.) Regions are counted if more than 10 measurements occurred in its bounding box.

From the overlay plot in Figure 9a we determine that the device covered all of the navigable floor space in the room over the six collection periods. Some

locations in the floor plan appear to have been missed by all collections, but these locations were not navigable in practice because they were blocked by temporary objects (e.g. boxes, equipment.) While an significant portion of the space was only covered by one period, it's unlikely to be due to some physical obstruction or complexity in the space. Given that four of six collection periods sampled the remote, lower-left nook, it's unlikely that the host did not have the opportunity to explore this region; rather, the variance in sample consistency is probably due to the nuances of the vacuum's navigation method.

Only 2.75 m^2—a relatively small portion of the space—was covered by every collection period. The area most represented by every collection period is nearest to the host's docking station. While the host has the capacity to broadly survey the majority of the space, it cannot do so in a single collection period. From Figure 9b we suggest that at least three surveys were necessary to visit a majority of the space. This room contains certain challenges, such as a desks, chairs and small corridors for the device to have to navigate around; simpler rooms may not have the same constraints.

CHAPTER V

RELATED WORKS

5.1 Mobile Sensing

The use of mobile devices to collect IEQ data has been explored previously with similar motivations. Work by Jin et al. (2018) proposed a mobile sensor that is capable of autonomously navigating throughout a room with a positioning system. The authors chose to use Simultaneous Location and Mapping (SLAM) implemented with a Microsoft Kinect. The mobile portion of the sensor was developed with a configurable robotic base, allowing sensors to be mounted at multiple heights. The device periodically polled the sensor data, associating it with approximate locations in space. The authors performed an experiment to evaluate the device's performance by activating it in a controlled space, gradually introducing CO_2 into the space, and measuring the result. Their experiment was controlled by placing fixed sensors at known locations in the space to verify the fidelity of sensor data and of the location data The authors use this experiment to argue for their *global trend* interpolation method that weights local values against global trends to refine to a more accurate result. Their experiment measuring CO_2 concentration validates this, but its application in other respects may less obvious.

Reggente et al. (2010) provide an analysis of mobile device called a *DustBot,* that navigates through public spaces, cleaning paths and collecting waste. The devices we equip with a sensor capable of measuring several properties including NO_2, O_3, CO, and PM concentrations. Their analysis demonstrates the unique capability of a mobile sensor to directly measure the source of a environmental attribute where fixed sensors may otherwise struggle. The authors model their data

with a gaseous distribution kernel Lilienthal, Reggente, Trincavelli, Blanco, and Gonzalez (2009).

5.2 Human-Aided Sensing

Research by Ulaganathan, Read, Collins, and Vincent (2017, 2019) preformed mobile IEQ data collection using sensors attached to human participants. The participants wore a Phillips Actiwatch 2, which periodically recorded illuminance over several days. This approach raises the sensor to a human scale, allowing the authors to associate their measurements directly with human activities. In practice, they consider measurements greater than 1 000 lx to represent outside daylight exposure. Where the goal human health and comfort is concerned, this approach eliminates the building as a constraint to collecting data.

CHAPTER VI

DISCUSSION

Our analysis demonstrates that a single mobile sensor reliably can collect IEQ data; further we have shown that the IEQ data we collected can be used to representatively characterize a space. This is done with significantly fewer expenses and minimal alterations to the existing built environment. The balance between utility and minimal interference contributes to the overall viability of our design. This chapter discusses how we utilized the mobile sensor, as well as the limitations of our methods and any future work in this area.

6.1 Utilization

The study occurred over a relatively short duration of five days. To optimally evaluate the mobile sensor we operated the system on a frequency that is not typical for most autonomous vacuum cleaners. The vacuum's existing programming only even allows for a single hour long cleaning cycle once per day. However, our method utilizes the non-standard operation frequency we deployed fairly heavily. This is a trade-off that we are faced in this study in favor of demonstrating the viability and reliability of the sensor system.

The mobile sensor was robust enough that we largely remained absent while the device operated over the five days. Access to the room was restricted, but the space is shared by several other people. We cannot be sure that other occupants did not interrupt the device, which threatens the validity of our study. Coexistence with occupants is a common property of indoor spaces, representing challenges that a mobile sensor would face in the real-world. Further, properties in the data would emerge if the device was interfered with; physically moving the host would

appear in the location estimation network's results. We did not find any evidence to suggest other occupants interfered with the device.

6.1.1 Design Objectives 1: Improved Spatial Resolution.

Through the analysis of the results we have demonstrated that the mobile sensor system collects measurements that carry a finer spatial resolution than traditional fixed sensors. This results in the capability to directly observe certain phenomena in the space (e.g. heat sources) instead of assuming their behavior. We show that our results in aggregate behave similar to fixed sensor data while only needing a single sensor.

6.1.2 Objective 2: Minimized Interference. The use of robotic vacuum cleaner allows the mobile sensor to largely "drop-in" to any environment that already supports such devices. While the mobile sensor makes the vacuum taller by approximately 10 cm, there are few other changes necessary for existing environments to support the mobile sensor system. In this implementation we utilized UWB sensors for their notable accuracy, however there may be more accessible location estimation technologies that we have overlooked. For example, WIFI-based positioning systems are similarly well established, but typically yield less accurate results Ma, Guo, Hu, and Xue (2015); Yang and Shao (2015).

6.1.3 Objective 3: Flexible Deployment. At the beginning of the study we required approximately thirty minutes to assemble the components, select a location of the host's docking station, as well as locations for the UWB anchors. Two additional steps were required. First, setting the `crontab` on the single-board computer before to our desired interval. Second, programmed into each of the UWB anchors, a relative position in space which it will occupy. This task is a This is a small inconvenience, but it does represent an opportunity to introduce

Table 3. Hardware expenses.

Item	UWB radio	Raspberry Pi	BME280	TSL2561	Create 2 roomba	Cell phone battery
Amount	4	1	1	1	1	2
Unit Price	$20	$30	$15	$15	$200	$10

human error; however, anchor self-localization methods such as Shi, Zhao, Cui, Lu, and Jia (2019) address this issue. At this point, the device was complete and fully functional. The device could feasibly be deployed in any similar indoor space with relative ease. There are limits to the size of the space the system will work in. The UWB sensors have a limited range of approximately 100 m, which would rule out excessively large spaces. Similarly, the host has a fixed battery life and cannot navigate large rooms effectively. The study largely demonstrates that the mobile sensor system is suitable for under-served locations similar to the room used in the study.

6.1.4 Expenses. The hardware costs for each of the components in the mobile sensor system total approximately $350 USD. A cost breakdown is provided in Table 3. A promotional credit was used to cover the expenses of Google Cloud Provider and an additional $10 USD was used to purchase the plastic material and bonding agents used to create the bumper extension. Comparably, the Onset HOBO sensors used in the study typically cost over $600 USD.

Cost is relevant here to both our practical capacity to perform research as well as to the capability of this system to be utilized in under-served buildings. Conventional systems often depend upon fairly expensive building systems, while

our mobile sensor is reliant on a common robotic vacuum and a location estimation network. The UWB location estimation network is likely the most uncommon hardware component of the sensor system; however, they are fairly inexpensive and UWB radios for other tasks are becoming common in consumer electronics.

6.2 Limitations

This sections discusses the limitations to the software components, physical design, and general architecture that we have identified in our mobile sensor throughout the pilot study.

6.2.1 Host System Implementation. The host software was written such that the position, the available measured values, and timestamp are paired together into a spatiotemporal value. By doing this, we are blocking our the procedure, waiting on several separate hardware components to achieve the spatiotemporal value. This adds inherent error—even if it is very minor—because the values, though paired, may have been read out at slightly different times. In practice this error is small, on the order of single milliseconds. Separating each of these values into their own time-series, and discretizing into spatiotemporal measurements post-hoc may provide a mechanism for mitigating this error. At our scale with only a few sensor read-outs, this is mostly inconsequential; however, a system that aims to collate a greater number of sensor read-outs see this error accumulate.

6.2.2 Sensor Height. In the study we fixed the mobile sensor height, as well as the fixed HOBO sensors at a height of 15 cm above the floor. We chose this height because we wanted the device to be able to explore below furniture objects as much as possible. For other applications, measurements at typical waist height (300 cm), head height (600 cm), or ceiling height (800 cm) may be more

desirable than floor height. For example, researchers interested in the effects of daylight on productively most are likely interested in the illuminance at head height.

6.3 Conclusion

This thesis gave an analysis of a mobile sensor, capable of measuring IEQ data with a fine spatiotemporal resolution. This analysis has demonstrated that the mobile sensor: (i) reliably collects spatiotemporal measurements, (ii) requires minimal modification to the existing building environment, and (iii) is suitable for environments that are traditionally under-served by existing sensor systems. These performance criteria were evaluated through a pilot study, wherein the sensor system collected approximately 100 000 measurements in ten hours over a five day interval. The implementation of the system is detailed, along with the computational methods used to evaluate the results.

APPENDIX

THE FIRST APPENDIX

Below are time-series plots of illuminance and temperature data taken on 8 January 2021; illuminance has been converted to log-space and illuminance values less than 1 are clamped to 0. There is a greater degree of variance the illuminance plots, given the natural variability of daylight. Discontinuities are visible in both plots, though more easily identified in the temperature plots. Both plots demonstrate consistent patterns consistent with daylight/heating behaviors. Low variance in the illuminance data is notable here as it may represent the effects of electric lighting absent daylight.

Figure A.10. Illuminance (log-space) time-series plots

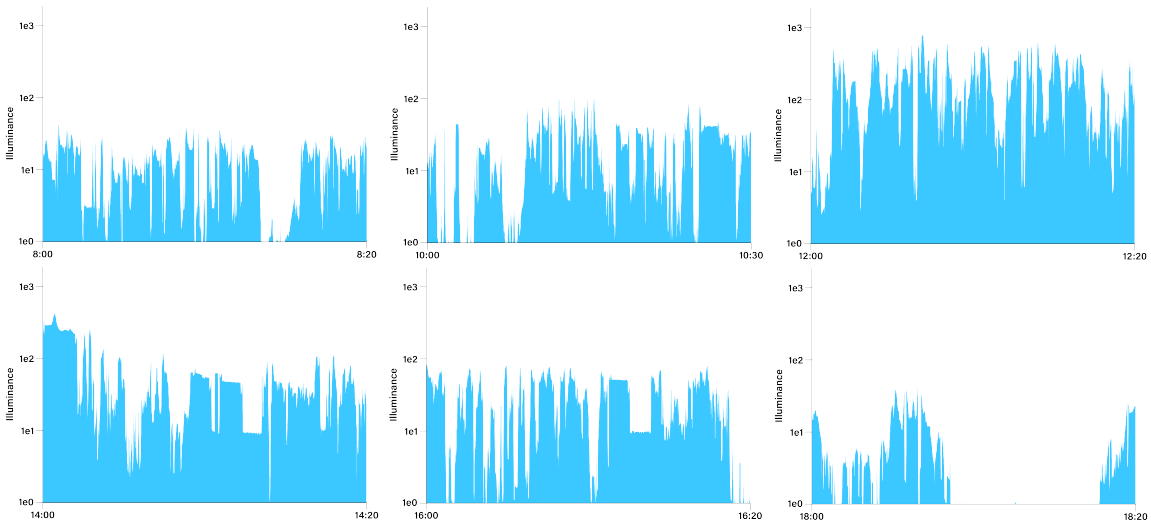
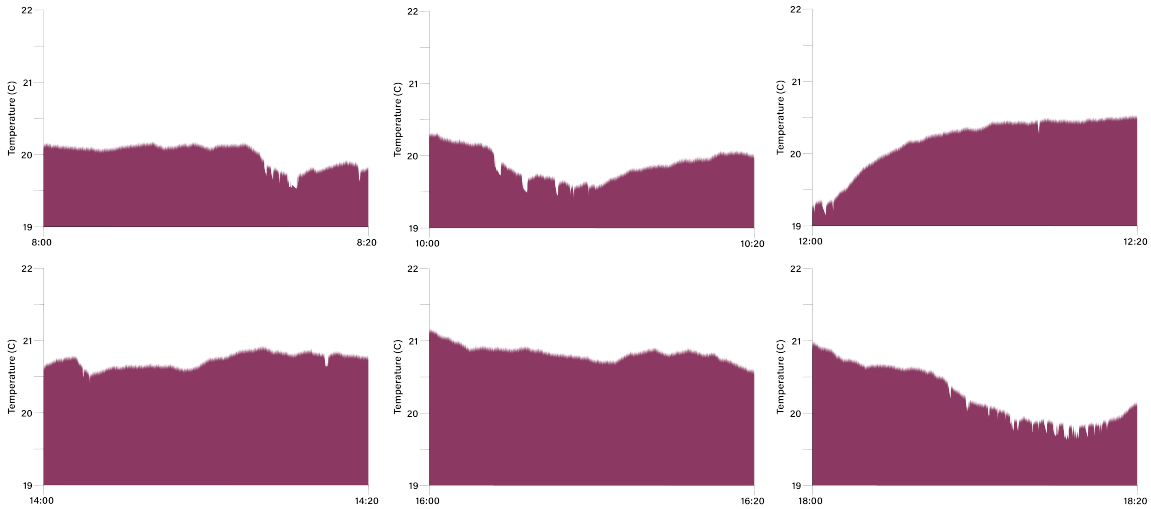


Figure A.11. Temperature time-series plots



REFERENCES CITED

Amundadottir, M. L., Rockcastle, S., Khanie, M. S., & Andersen, M. (2017). A human-centric approach to assess daylight in buildings for non-visual health potential, visual interest and gaze behavior. *Building and Environment*, *113*, 5–21.

Baccarelli, A., Barretta, F., Dou, C., Zhang, X., McCracken, J. P., Díaz, A., ... Hou, L. (2011). Effects of particulate air pollution on blood pressure in a highly exposed population in beijing, china: a repeated-measure study. *Environmental health*, *10*(1), 1–10.

Bluyssen, P. M. (2013). *The healthy indoor environment: How to assess occupants' wellbeing in buildings*. Routledge.

Castillo-Effer, M., Quintela, D. H., Moreno, W., Jordan, R., & Westhoff, W. (2004). Wireless sensor networks for flash-flood alerting. In *Proceedings of the fifth ieee international caracas conference on devices, circuits and systems, 2004.* (Vol. 1, pp. 142–146).

Cree, J. V., Dansu, A., Fuhr, P., Lanzisera, S. M., McIntyre, T., Muehleisen, R. T., ... Castello, C. (2013). *Sensor characteristics reference guide* (Tech. Rep.). Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Decawave. (2020). Real time location systems: An introduction.

(`https://www.decawave.com/sites/default/files/resources/aps003_dw1000_rtls_introduction.pdf` Accessed 2020-01-10)

Falsi, C., Dardari, D., Mucchi, L., & Win, M. Z. (2006). Time of arrival estimation for uwb localizers in realistic environments. *EURASIP Journal on Advances in Signal Processing*, *2006*, 1–13.

Fernandez, N. E., Katipamula, S., Wang, W., Xie, Y., Zhao, M., & Corbin, C. D. (2017). *Impacts of commercial building controls on energy savings and peak load reduction* (Tech. Rep.). Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... Oliphant, T. E. (2020, September). Array programming with NumPy. *Nature*, *585*(7825), 357–362. Retrieved from `https://doi.org/10.1038/s41586-020-2649-2` doi: 10.1038/s41586-020-2649-2

Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. doi: 10.1109/MCSE.2007.55

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Jin, M., Liu, S., Schiavon, S., & Spanos, C. (2018). Automated mobile sensing: Towards high-granularity agile indoor environmental quality monitoring. *Building and Environment*, *127*, 268–276.

Lan, L., Lian, Z., & Pan, L. (2010). The effects of air temperature on office workers' well-being, workload and productivity-evaluated with subjective ratings. *Applied ergonomics*, *42*(1), 29–36.

Li, Y., Zhuang, Y., Zhang, P., Lan, H., Niu, X., & El-Sheimy, N. (2017). An improved inertial/wifi/magnetic fusion structure for indoor navigation. *Information Fusion*, *34*, 101–119.

Lilienthal, A. J., Reggente, M., Trincavelli, M., Blanco, J. L., & Gonzalez, J. (2009). A statistical approach to gas distribution modelling with mobile robots - the kernel dm+v algorithm. In *2009 ieee/rsj international conference on intelligent robots and systems* (p. 570-576). doi: 10.1109/IROS.2009.5354304

Ma, R., Guo, Q., Hu, C., & Xue, J. (2015). An improved wifi indoor positioning algorithm by weighted fusion. *Sensors*, *15*(9), 21824–21843.

Mainwaring, A., Culler, D., Polastre, J., Szewczyk, R., & Anderson, J. (2002). Wireless sensor networks for habitat monitoring. In *Proceedings of the 1st acm international workshop on wireless sensor networks and applications* (pp. 88–97).

Nezis, I., Biskos, G., Eleftheriadis, K., & Kalantzi, O.-I. (2019). Particulate matter and health effects in offices-a review. *Building and Environment*, *156*, 62–73.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Reggente, M., Mondini, A., Ferri, G., Mazzolai, B., Manzi, A., Gabelletti, M., ... Lilienthal, A. J. (2010). The dustbot system: Using mobile robots to monitor pollution in pedestrian area. *Proc. of NOSE*.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, *20*, 53–65.

Shi, Q., Zhao, S., Cui, X., Lu, M., & Jia, M. (2019). Anchor self-localization algorithm based on uwb ranging and inertial measurements. *Tsinghua Science and Technology*, *24*(6), 728–737.

Ulaganathan, S., Read, S. A., Collins, M. J., & Vincent, S. J. (2017). Measurement duration and frequency impact objective light exposure measures. *Optometry and Vision Science*, *94*(5), 588–597.

Ulaganathan, S., Read, S. A., Collins, M. J., & Vincent, S. J. (2019). Influence of seasons upon personal light exposure and longitudinal axial length changes in young adults. *Acta ophthalmologica*, *97*(2), e256–e265.

U.S. Department of Energy. (2011). *Energysavers: Tips on saving money energy at home, prepared by the office of energy efficiency renewable energy* (Tech. Rep.). U.S. Department of Energy, Washington, U.S. Retrieved from `https://www.eia.gov/outlooks/aeo/`

U.S. Energy Information Administration. (2020). *Annual energy outlook* (Tech. Rep.). U.S. Energy Information Administration, Washington, U.S. Retrieved from `https://www.eia.gov/outlooks/aeo/`

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. doi: 10.1038/s41592-019-0686-2

Yang, C., & Shao, H.-R. (2015). Wifi-based indoor positioning. *IEEE Communications Magazine*, *53*(3), 150–157.

Zafari, F., Gkelias, A., & Leung, K. K. (2019). A survey of indoor localization systems and technologies. *IEEE Communications Surveys & Tutorials*, *21*(3), 2568–2599.