Exploring Effective Sensing Indicators of Loneliness For Elderly Community in US and Japan

Xiayan Ji
University of Pennsylvania
United States
xjiae@seas.upenn.edu

Ahhyun Yuh
University of Pennsylvania
United States
lucy0828@seas.upenn.edu

Viktor Erdélyi
Osaka University
Japan
viktor@ist.osaka-u.ac.jp

Teruhiro Mizumoto
Osaka University
Japan
Chiba Institute of Technology
Japan
teruhiro.mizumoto@p.chibakoudai.jp

Hyonyoung Choi
University of Pennsylvania
United States
hyonchoi@seas.upenn.edu

Sean Lee Harrison
University of Pennsylvania
United States
seanha@nursing.upenn.edu

Emma Cho
University of Pennsylvania
United States
emmacho@upenn.edu

James Weimer
Vanderbilt University
United States
james.weimer@vanderbilt.edu

Hajime Nagahara
Osaka University
Japan
nagahara@ids.osaka-u.ac.jp

Teruo Higashino
Kyoto Tachibana University
Japan
Osaka University
Japan
higashino-t@tachibana-u.ac.jp

George Demiris
University of Pennsylvania
United States
gdemiris@upenn.edu

Oleg Sokolsky
University of Pennsylvania
United States
sokolsky@cis.upenn.edu

Insup Lee
University of Pennsylvania
United States
lee@cis.upenn.edu

ABSTRACT

Loneliness has long affected the elderly community. This issue is significantly worsened by the social isolation resulting from the COVID-19 pandemic. To address this pressing issue, we employed a sensor-based methodology to predict loneliness and potentially inform interventions. We deployed sensors in the residences of 22 elderly participants from US and Japan, gathering daily activities data through 22 sensor features. Given the extensive feature set, we identify the most effective sensors to ensure unobtrusiveness while upholding privacy. Regression analysis of these features revealed that our best-performing Random Forest model achieved an R^2 value of 0.86, on par with existing literature. In addition, we found that the sleep mattress sensor and temperature-humidity sensor were particularly indicative of loneliness. In summary, our research contributes to the HCI literature with effective non-invasive sensing modalities in assessing elderly loneliness, together with insights from our real-world sensor deployments in US and Japan-based elderly communities.

CHI EA '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0331-7/24/05.
https://doi.org/10.1145/3613905.3637123

ACM Reference Format:
Xiayan Ji, Ahhyun Yuh, Viktor Erdélyi, Teruhiro Mizumoto, Hyonyoung Choi, Sean Lee Harrison, Emma Cho, James Weimer, Hajime Nagahara, Teruo Higashino, George Demiris, Oleg Sokolsky, and Insup Lee. 2024. Exploring Effective Sensing Indicators of Loneliness For Elderly Community in US and Japan. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3613905.3637123

1 INTRODUCTION

The COVID-19 pandemic has exacerbated elderly loneliness and social isolation, leading to significant deterioration in their mental and physical well-being. This not only affects individuals but also poses challenges for healthcare systems and societies globally [8, 36, 38, 42]. To address this issue quantitatively, HCI researchers have turned to various sensing modalities to detect loneliness in the elderly [9, 15]. Sensor data makes it possible to build predictive models of loneliness, paving the way for timely interventions. However, introducing multiple sensors requires careful practical considerations, especially given the varying degrees of tech-familiarity among the elderly. This study aims to streamline this process by pinpointing the most effective sensing modalities for predicting loneliness, allowing for optimized sensor deployment, intelligent design and placement. The goal is to ensure a less intrusive and privacy-respecting experience for the elderly.

Our study is conducted in two distinct cultural contexts: the United States and Japan [23, 37]. We collect data from 22 participants across these regions over a period of five months, using a sensor suite composed of seven sensor types [22]. Our initial data analysis on user demographics data confirms that social isolation leads to increased feelings of loneliness among the elderly. From the sensor data, we extract 22 features that best capture the daily activities of the elderly. Our regression analysis of these features shows that the Random Forest model outperforms the other models tested, achieving an R^2 value of 0.86. Additionally, features from non-invasive sensors, namely the sleep mattress and temperature-humidity sensors, emerge as the most predictive, thereby significantly reducing the number of sensors required.

Additionally, our research sheds light on the challenges and opportunities of elderly technology adoption. We share deployment experiences categorized by engineers, coordinators, and study team roles. Based on these insights, we recommend strategies to optimize sensor deployment and user experience for the elderly. Future HCI researchers can leverage this guidance when organizing teams for similar studies.

Our contributions are outlined as follows: **(1) Cross-Cultural Deployment**: We deployed a passive sensing system for seniors in both the US and Japan, demonstrating its unique adaptability to diverse cultural contexts. **(2) Key Sensors Identification:** We identify key sensors for detecting loneliness using regression analysis, drawing from combined data from the US and Japan; **(3) Elderly Interaction with Technology:** We provide insights into how the elderly interact with technology, emphasizing challenges in sensor maintenance and potential avenues to enhance usability. The subsequent sections of this paper encompass a literature review in Section 2, detailed sensor importance analysis in Section 3, distilled insights in Section 4.2, discussion on limitations and potential interventions in Section 5, and conclusions in Section 6.

2 LITERATURE REVIEW

Sensor Modalities in Elderly Loneliness Assessment. Sensor technology has recently been leveraged to assess elderly well-being, specifically regarding loneliness and social isolation over the past decade in European countries such as Sweden [7], Norway [41], and Denmark [47]. Motion sensors capture mobility patterns that reflect mental health and sociability nuances [19, 21, 24]. Sleep quality is gauged using devices like Fitbit and sleep mattress sensors [16]. Wearables, with step counters or gyroscopes, track physical movements, indicating activity levels [12]. Contact sensors unobtrusively monitor activities such as cooking and hygiene, hinting at potential social isolation [19]. Passive sensors, like home appliance and temperature-humidity sensors, provide further insights [43]. While numerous sensors have been deployed to gauge elderly loneliness, there is insufficient research comparing their effectiveness across various cultural settings [18]. In addition, the literature on the optimization of sensor deployment and the enhancement of usability for the elderly is relatively sparse.

Feature Importance Analysis. Feature importance techniques vary from model-dependent to model-agnostic methods [14]. While model-dependent methods are intrinsic to specific models, model-agnostic ones like Permutation Importance (PFI) [10, 17] and SHapley Additive exPlanations (SHAP) [35] offer universal applicability. These techniques can provide insights either globally, across all observations, or locally, for specific data points. Our study emphasizes global model-agnostic methods for broad relevance, leveraging regression model coefficients, PFI, and SHAP values for comprehensive feature significance evaluation.

3 METHODOLOGY

Upon obtaining IRB approval, we collect data from 22 participants (17 from the US and 5 from Japan) from April to September 2023, yielding 140 weeks of data on 22 features, paired with weekly UCLA loneliness scores [39]. This data is gathered from seven distinct sensor types: motion, contact, BLE tag-based proximity, temperature and humidity, power meter, sleep mattress, and activity trackers [2–4]. We first describe our pre-processing methods and the features extracted for each data type. Then, we provide a descriptive summary the data collected, together with preliminary correlation analysis. Finally, we discuss feature selection outcomes from three regression models.

3.1 Data Collection Procedure

This subsection delineates the procedure for data collection, detailing the setting, participant recruitment, and ethical protocols.

Settings. In the US's greater Philadelphia area, participants were independent seniors living in senior apartments. In Japan's Shikanodai community, they resided in private homes with diverse living arrangements. Data collection began with a baseline session where qualified research coordinator gathered demographic data and administered the Mini-Mental State Examination (MMSE) to participants. After obtaining consent, tablets were provided for completing baseline and weekly surveys (Figure 4). Participants were strongly encouraged to complete the weekly 20-item UCLA loneliness survey. Concurrently, the study team installed sensor suites while participants answered baseline questions.

Recruitment. Recruitment was facilitated through community sessions, leading to the voluntary participation of 17 individuals from the U.S. with no dropouts noted to date of this work. In Japan, six community leaders have been recruited, with efforts to enlist more underway. The evaluation of sensor data validity is in progress and will be reported in the full data analysis.

Ethical Considerations Ethical compliance was ensured by a coordinator who thoroughly explained the consent form to participants in a baseline session before any sensor installations were carried out.

3.2 Sensing Feature Processing

Motion and Contact Sensor. For these sensors, we initially extract records indicating motion detection or an open contact sensor status. These records are sorted by timestamp, and only those separated by at least five minutes are retained to minimize redundant event reporting. Subsequently, daily movement or door opening occurrences in specific areas (entrance, kitchen, bathroom) are determined by counting daily records from the respective sensors, which yields six features in total from motion and contact sensors.

BLE Tag-Based Proximity Sensor. For our BLE tag-based sensor, we smooth the RSSI signal using a moving average over 300 records, then classify each timestamp as outdoor, indoor, or unsure. RSSI signals stronger than -55 dBm are labeled as indoor, while those weaker than -88 dBm are categorized as outdoor, based on established thresholds [32]. Signals falling between these thresholds are labeled as "unsure", potentially indicating transitional moments of entering or exiting the apartment. A missing RSSI signal suggests the participant has left the apartment.

Continuous outdoor readings are identified from consecutive unsure-to-outdoor transitions. We count daily outing events, excluding days with more than ten events to filter anomalous readings. Durations of outings are determined from consecutive outdoor-classified records. Only outings longer than a minute are considered. We then match outing events with durations using timestamps, yielding the features outing frequency ('tag_cnt') and event duration ('tag_duration').

Temperature and humidity sensor. For these sensors, we calculate weekly averages of the readings. Additionally, we identify concurrent peaks in temperature and humidity to detect showering events, which are characterized by peaks that occur no more than two minutes apart. In total, three features are extracted: mean temperature, mean humidity, and the number of showering events ('show_cnt').

Power meter sensor. For the smart power plug, which measures power usage, we connect it to an elderly individual's television. Given the occasional fluctuations in the TV's power consumption, we employ the K-means algorithm [27] to segment the power readings into two clusters. Among these, the cluster with the higher centroid value is interpreted as 'TV on', while the other is treated as 'TV off'. We then determine the 'daily total TV hours' feature ('tv') by totaling the number of 'TV on' instances within a day.

Sleep mattress sensor. We choose seven key metrics from standard sleep medicine literature out of a possible 27 variables [25]. Additionally, we incorporated two nap metrics by counting naps recorded before 6 PM. The nine metrics we utilized are defined as follows:

- REM Episodes: This measures the number of REM sleep phases.
- Sleep Efficiency: The ratio of total main sleep time to the time spent in bed.
- Sleep Latency: The time it takes to fall asleep once in bed.
- Total Sleep Time: The cumulative time spent during main sleep.
- Total Time in Bed: The entire duration spent in bed during main sleep.
- Wake-up Latency: The time taken after waking up before leaving the bed.
- WASO (Wakefulness After Sleep Onset): The time spent awake in bed after initially falling asleep for the night.
- Nap Time in Bed: The entire duration spent in bed during nap.
- Nap Sleep Time: The cumulative time spent during nap.

We calculate the weekly mean of these sleep metrics, omitting days with missing data.

Activity tracker. We distributed activity trackers to elderly participants, offering them the option to wear these devices. From these trackers, we collected daily step count data and calculated the weekly average, excluding any days when the tracker was not worn by the participants.

4 RESULTS

4.1 Quantitative Results: Data Analysis

Descriptive Statistics Table 1 summarizes the statistics of loneliness and social isolation. The average UCLA Loneliness score for the U.S. cohort is 46.41, indicating moderate loneliness with considerable variance (standard deviation 14.83). The mean Lubben Social Network Scale (LSNS) is 15.47, suggesting low social isolation risk, though four participants score below 12, indicating higher risk. The Japanese cohort shows milder loneliness (average UCLA 40.40, standard deviation 6.66) and no significant social isolation risk (minimum LSNS 12). With only four data points from Japan, we used Cohen's d (Table 3) to compare US and Japan datasets. All sensor features show absolute d values below 0.8, indicating no significant differences, allowing us to combine both datasets for further analysis.

Our overview of 22 sensor features (Tables 4, 5) reveals distinct patterns: movement events are more frequent than door openings. On average, entrance door interactions occur five times daily, while motion sensors record about 16.88 movements per day in the same area, likely influenced by pet activity. Tag sensor data shows the elderly leave their homes 3.41 times daily for an average of 41.86 minutes. The daily step count varies widely from zero to 7783, averaging 1039.64 steps, with a median of 505.71 indicating sporadic high-activity days. Sleep data indicates an average of 6.47 hours sleep per night with 89% efficiency, 20 minutes to fall asleep, and 39 minutes to the first wake episode. Napping varies greatly among participants. Demographically (Table 2), our sample has an equal number of male and female participants, mostly retired, with education levels mostly at or above bachelor's, including six with advanced degrees.

Baseline Correlation Analysis Our analysis begins with the Pearson's (P) and Spearman's (S) correlations [31] assessment of the baseline survey data, which includes demographics, the LSNS, and blood pressure metrics. Pearson's correlation measures the linear relationship between two variables, while Spearman's correlation assesses the monotonic relationship based on their ranks [40]. The result is summarized in Table 2. Notably, LSNS negatively correlates with UCLA in both US and Japanese samples. The finding suggests an increased social isolation often amplifies feelings of loneliness. In the US cohort, gender appears to influence loneliness, with male

	UCLA		LSNS	
	US	Japan	US	Japan
mean	**46.41**	**40.40**	15.47	17.60
50%	46.50	43.00	14.00	15.00
std	**14.83**	**6.66**	7.76	6.19
min	20.00	30.00	**3.00**	**12.00**
max	**73.00**	47.00	27.00	28.00

Table 1: UCLA Loneliness Scale (UCLA) and Lubben Social Network Scale (LSNS) Statistics

	US Baseline		Japan Baseline	
	P	S	P	S
Age	0.17	0.26	-0.30	-0.05
Gender	**0.38**	**0.36**	-	-
Education years	-0.02	0.01	**-0.82**	-0.22
Job status	-0.14	-0.10	-	-
LSNS	**-0.73**	**-0.70**	**-0.91**	**-0.67**
Diastolic blood pressure	0.05	0.14	0.29	0.20
Systolic blood pressure	0.10	-0.05	**-0.54**	**-0.56**

Table 2: Pearson's (P) and Spearman's (S) correlation coefficients between the baseline survey data and UCLA, the relatively significant correlations are in bold.

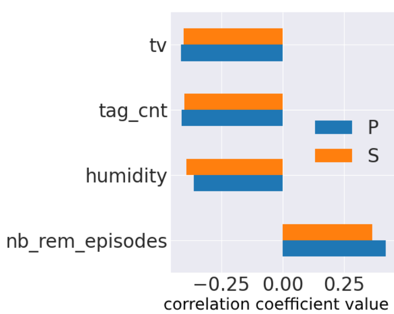


Figure 1: Top Pearson's and Spearman's correlation coefficients between sensor features and the UCLA loneliness scores.

Demographic	Total (N)	Percentage (%)	Male	Female
Age Group:				
65-70	7	31.82%	3	4
71-76	8	36.36%	2	6
77 or above	7	31.82%	6	1
Education Years:				
Bachelor's	16	72.73%	6	10
Master's and beyond	6	27.27%	5	1
Occupation:				
Retired	21	95.45%	10	11
Other	1	4.55%	1	0

Figure 2: Participant demographics summary.

participants exhibiting an average UCLA 21.84% higher than that of the female participants. In the Japanese sample, education years and systolic blood pressure may be inversely related to loneliness.

Sensor Correlation Analysis We evaluated the Pearsons'(P) and Spearman's (S) correlation between weekly sensor averages (Figure 3) and weekly UCLA loneliness scores (Figure 1) similar to the baseline correlation analysis. The inter-sensor correlation assessment aids in addressing multicollinearity and potentially reducing feature count. Assessing sensor-loneliness correlations provides a foundational insight for subsequent regression analysis. Only relatively significant correlations are depicted for clarity in Figure 3 and Figure 1.

Figure 3 reveals an evident correlation between total nap time in bed and sleep efficiency (P = 0.98, S = 0.94), step counts and total nap time (P = 0.95, S = 0.96), implying that more activity might lead to longer rest periods. Movements in the bathroom and near the entrance are related (P = 0.72, S = 0.66), possibly due to common senior residence layouts. Humidity and temperature, both measured in the bathroom, are correlated (P = 0.84). Interestingly, kitchen contact sensor activity positively correlates with total sleep time (P = 0.71) and bed duration (P = 0.70), suggesting that more kitchen activity may lead to extended sleep. Researchers might consider excluding activity trackers and kitchen contact sensors, given their high correlation with sleep mattress sensor metrics.

From Figure 1, we observe that REM episode counts, humidity levels, tag counts, and TV durations emerge as better loneliness indicators than the rest, each with correlations exceeding 0.35. Notably, a rise in REM episodes points to increased feelings of loneliness [20, 44], which could be caused by sleep disorders and lower sleep quality [33, 48]. On the other hand, higher humidity levels, which might suggest regular showers, and more outdoor activity, as tracked by tag counts, are tied to lower feelings of loneliness. Interestingly, extended TV hours negatively correlate with loneliness, suggesting entertainment might counter such feelings in our elderly cohort.

Regression Models and Feature Selection To predict weekly loneliness levels, we employ three regression models: Linear Regression [28], Random Forest [29], and Elastic Net Regression [26], chosen for their efficacy in feature selection. We perform five-fold cross-validation with sensor features as predictors and the weekly UCLA loneliness scores as the response. The results are assessed via metrics like coefficient of determination (R^2), Mean Absolute Error (MAE), Explained Variance (EV), and Mean Squared Error (MSE), as shown in Figure 6. Notably, the Random Forest model, with an R^2 of 0.86, matches state-of-the-art (SOTA) results from previous UCLA loneliness score studies [6, 46]. Nevertheless, we acknowledge the larger MSE observed in our study, likely attributable to the extensive range and variance of the target variable in our dataset, which is consistent with the known sensitivity of MSE to the scale of the data.

From our five-fold cross-validation, we highlight the best-performing fold, using regression model coefficients, permutation importance (PFI) [5], and SHapley Additive exPlanations (SHAP) [35] to gauge feature importance. For the Random Forest model, we refer to the Gini importance (Gini) [34], represented as the 'feature importance' in the sklearn library [1]. Features were validated for statistical significance using a p-value threshold smaller than 0.05.

Following the previous sensor correlation analysis, we filter out less significant sensor features during feature selection, especially

Feature	Value	Feature	Value	Feature	Value	Feature	Value
kitchen_movement	0.03	tv	-0.02	waso	-0.35	sleep_latency	-0.62
bathroom_movement	0.00	temperature	0.02	sleep_efficiency	0.18	wakeup_latency	-0.72
entrance_movement	-0.01	humidity	0.07	total_sleep_time	0.19	nb_rem_episodes	0.28
kitchen_contact	0.42	shower_cnt	0.01	nap_total_timeinbed	-0.55	total_timeinbed	-0.00
bathroom_contact	0.00	tag_cnt	-0.01	nap_total_sleep_time	-0.48	steps	-0.38
entrance_contact	0.04	tag_duration	-0.01				

Table 3: Cohen's d values for feature difference between US and Japan data. All features have absolute values smaller than 0.8, indicating no significant difference.

	Motion Sensor			Contact Sensor			Tag		TV	Temperature	Humidity	Shower
	Kitchen	Bathroom	Entrance	Kitchen	Bathroom	Entrance	Count	Duration	Hours			Count
mean	15.01	10.51	16.88	4.44	1.46	5.00	3.41	41.86	9.31	77.44	55.50	0.32
50%	15.00	10.00	16.00	2.00	0.00	4.00	3.00	4.00	7.13	77.61	55.10	0.00
std	7.99	6.22	9.35	5.93	2.52	4.19	2.28	77.77	8.29	3.06	9.94	0.72
min	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.04	0.00	65.98	30.14	0.00
max	42.00	38.00	50.00	37.0	15.0	20.0	9.00	355.06	34.98	88.95	85.40	4.00

Table 4: Basic sensor statistics.

	Steps	WASO (h)	Sleep Latency (h)	Wakeup Latency (h)	REM Episodes	Total Time in Bed (h)	Sleep Efficiency	Total Sleep Time (h)	Nap Time (h) in Bed (h)	Nap Sleep (h) Time (h)
mean	1039.64	0.65	0.33	0.03	2.82	7.21	0.89	6.47	1.10	0.77
50%	505.71	0.58	0.30	0.02	2.57	7.10	0.90	6.52	1.28	0.82
std	1174.31	0.48	0.16	0.04	1.51	0.87	0.06	0.88	0.81	0.66
min	0.00	0.00	0.13	0.00	0.00	0.50	0.48	0.28	0.00	0.00
max	7783.00	3.71	1.77	0.71	8.00	12.48	0.97	9.23	6.88	4.47

Table 5: Basic sensor statistics continued.

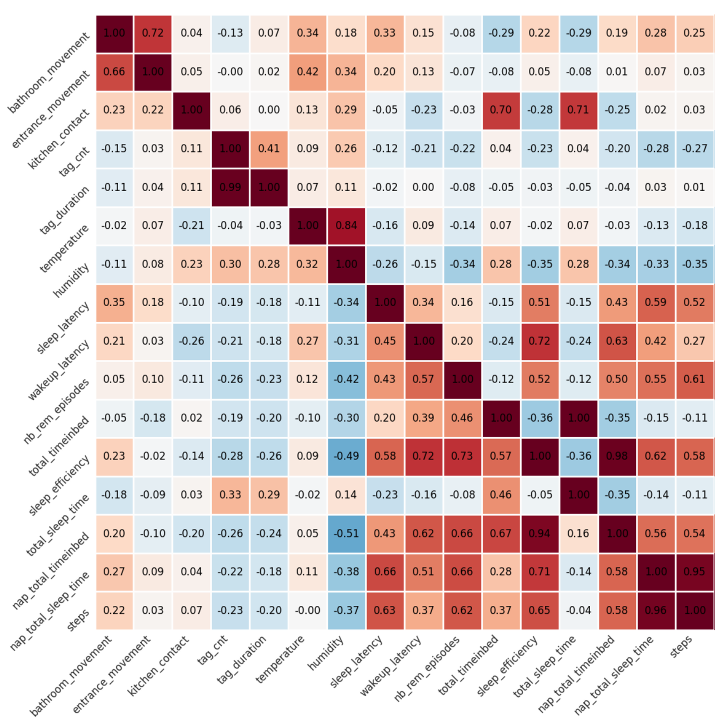
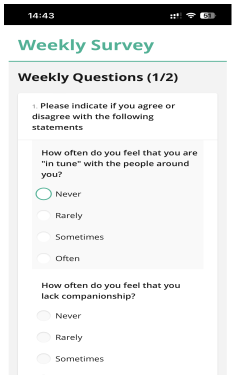
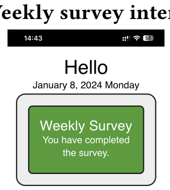


Figure 3: Heatmap for sensor correlations, where upper triangle is Pearson's correlation coefficient and lower triangle is Spearman's correlation coefficient.

(a) Weekly survey before completion.



(b) Weekly survey interface.



(c) Weekly survey after completion.

Figure 4: Screenshots for iCareLoop tablet app

when they correlate with more important features. Figure 5 depicts these results from the top-performing Random Forest model, with p-values annotated on each bar. Sleep mattress sensor metrics, like sleep efficiency and total sleep time, emerged as the most important. Shower count and humidity, from the temperature-humidity sensor, and door opening frequencies, from the entrance contact sensor, also appears to be valuable in loneliness prediction. Interestingly, all these sensors are non-invasive in nature.

In sum, our feature selection results derived from regression models consistently corroborate with our sensor correlation analysis, underscoring the predictive value of sleep mattress sensors and temperature-humidity sensors for assessing loneliness. By solely utilizing these two sensors, as indicated by 'Random Forest Reduced' in Table 6, we achieve performance comparable to SOTA literature. On the other hand, BLE tag-based sensor and power meter sensor are not primary contributors in the regression analysis, the entrance contact sensor display notable importance instead. Moreover, our study indicates that non-invasive sensors can efficiently gauge elderly loneliness, outperforming activity trackers within our chosen sensor suite.

In future studies targeting the elderly, we recommend the use of a minimal number of non-invasive sensors to capture crucial data features, thereby minimizing intrusion. Based on our analysis, researchers might consider employing sensors such as the sleep

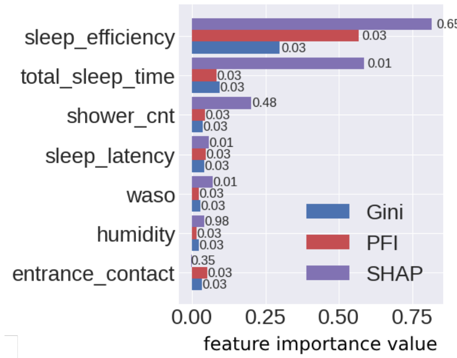


Figure 5: Feature importance of Random Forest model, with p-value annotated next to each bar.

Model	R^2	MAE	EV	MSE
Linear Model	0.67	5.71	0.73	54.79
Random Forest	**0.86**	3.70	**0.87**	25.22
Elastic Net	0.67	5.76	0.73	54.33
Baseline 1 [6]	0.35	**0.81**	-	**0.91**
Baseline 2 [46]	0.57	4.46	0.57	5.63
Random Forest Reduced	0.78	5.75	0.78	61.74

Figure 6: Regression model performance metrics, the best result are in bold.

mattress sensor and the temperature-humidity sensor. Our insights serve as a guide, assisting researchers in choosing the most relevant sensors that align with both the research objectives and the preferences of the elderly participants.

4.2 Qualitative Results: Lessons learned

In our study, we categorize team duties into roles: middleware, software engineers, research coordinators, and the study team. We discuss challenges and solutions for each role, aiming to offer role-specific insights for researchers working on elderly related studies in the HCI community.

4.2.1 Insights by Middleware Engineers. Middleware engineering covers various tasks, from sensor activation to ongoing device maintenance in elderly participants' residences.

Sensor Setup. The initial setup for mini PCs was lengthy and error-prone. We streamlined this by creating a module that replicates a disk image, reducing setup time from two hours to just 40 minutes. Additionally, manual entry of MAC addresses from multiple sensors pose accuracy issues. Our solution is an automated system that outputs the needed JSON configuration for the mini PC when fed a CSV with sensor details.

Sensor Installation. With multiple sensors per household, oversights were common. To address this, we implemented a double-confirmation checklist, requiring two installation engineers to cross-verify the setup of all devices, thereby ensuring a meticulous installation process. The diverse settings of elderly homes demand flexibility in installation. For example, older fridges with large gaps sometimes requires us to rethink contact sensor placements. As a solution, we install sensors on cupboards alternatively to monitor food-related activities, addressing the need for adaptability in real-world HCI scenarios.

Sensor Maintenance. Servicing a tech-heavy system with the elderly presents its own challenges, for example, we receive many reports on accidental unplugging of mini PCs due to cleaning. Our solution is to automate mini PCs for resilience and harnessing remote tools like TeamViewer, making them 'plug-and-run'. In addition, surface adherence of sensors, especially in moisture-rich environments like bathrooms and kitchens, was tricky. When optimal placements fails, we innovatively reposition sensors and document placements with photos, making troubleshooting more straightforward for both engineers and participants.

Our project emphasizes the use of non-intrusive sensors and requires minimal effort from the elderly. By sharing these insights, we aim to facilitate future HCI efforts by combining robust technical solutions with empathetic user engagement.

4.2.2 Insights by Software Engineers and ML Engineers. Software and ML engineers are central to the backend, developing tools and methodologies that inform the research. Their tasks include crafting intuitive software interfaces and refining data for valuable insights.

Survey Application Refinement. Designing for elderly users brought UI challenges to light. Feedback emphasized larger fonts for better readability. To cater to this, our engineers refined the UI, optimizing font sizes, button placements, color contrasts, and response mechanisms. This user-informed design ensured the tools were both accessible and tailored for our elderly participants.

Data Quality and Processing. Ensuring data quality, especially in ML-driven research, is paramount. During our data collection process, we encountered data inconsistencies and gaps. In response, we applied data imputation techniques and designed robust pipelines for diverse sensor types. Partnering with domain experts, our ML team tailored preprocessing and normalization methods. This endeavor accentuated the critical link between ML capabilities and data quality.

These experiences reinforce the interdisciplinary core of HCI. As we advance in technology, its accessibility, particularly for the elderly demographic, remains vital.

4.2.3 Insights by Research Coordinators. Research coordinators play a pivotal role from identifying potential participants to engaging them. Beyond recruitment, they are often the first to introduce the research agenda, setting its tone and impression.

Participant Recruitment. Recruiting from underserved older communities brought challenges, including monitoring fears, data misuse concerns, and a fundamental distrust of research. Direct outreach at senior facilities were often met with lukewarm responses, highlighting technological intricacy fears. To overcome this, we emphasized the study's non-invasive nature (activity tracker is optional), promoting transparency and simplifying outreach materials.

Technology Engagement. Participant feedback revealed varied technology comfort levels. While some participants easily adapted to tools like tablet questionnaires, others remained apprehensive. For example, user 2009 expressed his anxiety about using technology: *"I got a new router from Verizon and everything is screwed up ... I'm a tech-challenged old man. I'll try to get the tablet working, but I'm not promising anything."* Addressing this gap, we oriented participants about daily devices—activity trackers and tablets, on charging frequency and basic functionality. Concerns about monitoring were addressed by emphasizing the study's ethical foundation. Coordinators highlighted the research's intent to bolster elderly well-being rather than invade privacy. Clarifying data anonymization and the absence of audio or video sensors helped allay many concerns.

These insights demonstrate that engaging technology-averse demographics requires patience and tailored approaches. We hope our feature significance analysis will allow future studies to reduce the number of sensors, thereby minimizing sense of obtrusiveness. This addresses the elderly's concerns about monitoring while upholding their privacy.

4.2.4 Insights by Study Team. The study team plays a pivotal role in direct interactions with the elderly, encompassing activities such as data collection, periodic reminders for survey completion, and eliciting feedback.

Communicating with Participants. Engaging effectively with the elderly cohort presents its unique set of challenges, especially when establishing communication for baseline or troubleshooting sessions. Often, the elderly may inadvertently overlook phone calls or forget scheduled appointments. For instances, user 2001 mentioned: *"The bigger problem is I can't always find my phone. But I am trying to keep it close now."* To address this, we leverage a multimodal communication strategy that may involve text messages, voicemails, and traditional mails. Furthermore, to minimize the potential for missed appointments, we institute a protocol where participants receive a reminder call one day prior to any scheduled visit. To enhance our communication effectiveness and foster trust, we customize our interaction methods based on the preferences and comfort levels of the elderly. For instance, with user 2018, we introduced the sensors during a Bingo game session, leveraging a familiar and enjoyable setting to facilitate the introduction.

Survey Administration. Administering regular surveys to an elderly demographic comes with inherent challenges, primarily attributable to lapses in memory. While our initial approach favored daily surveys, the realization quickly dawned that this frequency, albeit comprising a concise three questions, posed a demand on the elderly. For instance, User 2004 says: *"Some time I got confused with my memory, not sure if I fill out the daily surveys and I need to constantly check the tablet to confirm."* In response to this challenge, we transitioned to advocating for the completion of end-of-week surveys. To ensure consistent data collection, we disseminate weekly text reminders, prompting participants to complete their surveys. For HCI researchers embarking on similar endeavors with elderly subjects, it may be prudent to consider a weekly survey cadence complemented by regular reminders.

Integrating Feedback. Feedback from our elderly participants underscored a desire for more tangible insights into the collected data. For example, User 2006 expressed: *"I would be interested in visual feedback of my data to help me better understand my activity levels."* Given this feedback, we anticipate that visualizing sleep mattress and temperature-humidity sensor features, delineated in previous sections, can pave the way for crafting intuitive and simple visualizations. Such visual representations could not only empower elderly participants by providing insights into their activity patterns, potentially aiding in recognizing feelings of loneliness, but could also serve as a valuable tool for caregivers to gather insights about the well-being of their beloved ones.

Our findings highlight the importance of a human-centric approach in HCI research with the elderly. Balancing data collection with comfort requires careful planning and incorporation of feedback.

5 LIMITATIONS AND DISCUSSION

A limited sample size in the Japanese dataset and a participant pool predominantly holding bachelor's degrees or higher may limit the findings of our study. Efforts to enlarge the sample size and diversify educational backgrounds are underway. These steps are crucial for deepening insights into the cultural nuances affecting participant interactions in the US and Japan. Additionally, methodological limitations include sensor variability, reliance on self-reported loneliness metrics, and inconsistencies in sensor installation and activity tracker usage. Data gaps due to participants' varied environments and potential inaccuracies in machine learning methods further complicate the analysis. These issues underscore the inherent complexities of real-world research, necessitating careful study design and execution. The exploration of technological solutions like digital companions and emotion recognition in addressing loneliness [13, 45], while promising [30], demands rigorous evaluation of effectiveness and ethical considerations [11].

6 CONCLUSION

In this study, we conduct a comprehensive evaluation of current sensing modalities for gerontological loneliness and identified the most effective categories. Using data from the US and Japan, we find that social isolation, intensified by COVID-19, has the potential to heighten feelings of loneliness. Our regression analysis stands on par with the state-of-the-art, with the Random Forest model achieving an R^2 of 0.86. We found that non-invasive sensors, specifically sleep mattress sensor and temperature-humidity sensor, serve as particularly informative indicators of loneliness. Prioritizing non-intrusiveness, our study highlights the importance of devising solutions customized for the elderly that include 'plug-and-run' system, survey reminders, and remote troubleshooting. From the standpoint of HCI research, embedding these design elements is crucial to optimize experiences and interventions targeting the aging demographic.

ACKNOWLEDGMENTS

This work was supported in part by a joint NSF-JST grant (NSF-2125561, JST-21453299) titled "SCC-IRG JST: Active sensing and personalized interventions for pandemic-induced social isolation".

REFERENCES

[1] 2021. RandomForestRegressor — scikit-learn 0.24.2 documentation. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html#sklearn.ensemble.RandomForestRegressor.feature_importances_ Accessed: 2023-10-11.

[2] 2023. SwitchBot. https://us.switch-bot.com/ Accessed: 10/11/2023.

[3] 2023. TileTag. https://www.tile.com/ Accessed: 10/11/2023.

[4] 2023. Withings. https://www.withings.com/ Accessed: 10/11/2023.

[5] André Altmann, Laura Toloşi, Oliver Sander, and Thomas Lengauer. 2010. Permutation importance: a corrected feature importance measure. *Bioinformatics* 26, 10 (2010), 1340–1347.

[6] Johanna Austin, Hiroko H Dodge, Thomas Riley, Peter G Jacobs, Stephen Thielke, and Jeffrey Kaye. 2016. A smart-home system to unobtrusively and continuously assess loneliness in older adults. *IEEE journal of translational engineering in health and medicine* 4 (2016), 1–11.

[7] Iman Azimi, Amir M Rahmani, Pasi Liljeberg, and Hannu Tenhunen. 2017. Internet of things for remote elderly monitoring: a study from user-centered perspective. *Journal of ambient intelligence and humanized computing* 8 (2017), 273–289.

[8] Jürgen Barth, Sarah Schneider, and Roland Von Känel. 2010. Lack of social support in the etiology and the prognosis of coronary heart disease: a systematic review and meta-analysis. *Psychosomatic medicine* 72, 3 (2010), 229–238.

[9] Taxiarchis Botsis, George Demiris, Steinar Pedersen, and Gunnar Hartvigsen. 2008. Home telecare technologies for the elderly. *Journal of Telemedicine and Telecare* 14, 7 (2008), 333–337.

[10] Leo Breiman. 2001. Random forests. *Machine learning* 45 (2001), 5–32.

[11] Johana Cabrera, M Soledad Loyola, Irene Magaña, and Rodrigo Rojas. 2023. Ethical dilemmas, mental health, artificial intelligence, and llm-based chatbots. In *International Work-Conference on Bioinformatics and Biomedical Engineering*. Springer, 313–326.

[12] Carly Cooper, Anne Gross, Chad Brinkman, Ryan Pope, Kelli Allen, Susan Hastings, Bard E Bogen, and Adam P Goode. 2018. The impact of wearable motion sensing technology on physical activity in older adults. *Experimental gerontology* 112 (2018), 9–19.

[13] Alan S Cowen, Petri Laukka, Hillary Anger Elfenbein, Runjing Liu, and Dacher Keltner. 2019. The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Nature human behaviour* 3, 4 (2019), 369–382.

[14] Arun Das and Paul Rad. 2020. Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey. arXiv:2006.11371 [cs.CV]

[15] George Demiris, Marilyn J Rantz, Myra A Aud, Karen D Marek, Harry W Tyrer, Marjorie Skubic, and Ali A Hussam. 2004. Older adults' attitudes towards and perceptions of 'smart home'technologies: a pilot study. *Medical informatics and the Internet in medicine* 29, 2 (2004), 87–94.

[16] Afsaneh Doryab, Daniella K Villalba, Prerna Chikersal, Janine M Dutcher, Michael Tumminia, Xinwen Liu, Sheldon Cohen, Kasey Creswell, Jennifer Mankoff, John D Creswell, et al. 2019. Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and fitbit data. *JMIR mHealth and uHealth* 7, 7 (2019), e13209.

[17] Aaron Fisher, Cynthia Rudin, and Francesca Dominici. 2019. All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. arXiv:1801.01489 [stat.ME]

[18] Hulya Gokalp and Malcolm Clarke. 2013. Monitoring activities of daily living of the elderly and the potential for its use in telecare and telehealth: a review. *TELEMEDICINE and e-HEALTH* 19, 12 (2013), 910–923.

[19] Nadee Goonawardene, XiaoPing Toh, and Hwee-Pink Tan. 2017. Sensor-driven detection of social isolation in community-dwelling elderly. In *Human Aspects of IT for the Aged Population. Applications, Services and Contexts: Third International Conference, ITAP 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II 3*. Springer, 378–392.

[20] Ramon Greenberg, Chester Pearlman, Wynn R Schwartz, and Hildreth Y Grossman. 1983. Memory, emotion, and REM sleep. *Journal of abnormal psychology* 92, 3 (1983), 378.

[21] Xiayan Ji, Xian Li, Ahhyun Yuh, Amanda Watson, Claire Kendell, James Weimer, Hajime Nagahara, Teruo Higashino, Teruhiro Mizumoto, Viktor Erdélyi, et al. 2023. Short: Integrated Sensing Platform for Detecting Social Isolation and Loneliness In the Elderly Community. In *2023 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. IEEE, 148–152.

[22] Xiayan Ji, Ahhyun Yuh, Hyonyoung Choi, Amanda Watson, Claire Kendell, Xian Li, James Weimer, Hajime Nagahara, Teruo Higashino, Teruhiro Mizumoto, et al. 2023. iCareLoop: Closed-Loop Sensing and Intervention for Gerontological Social Isolation and Loneliness. In *Proceedings of the ACM/IEEE 14th International Conference on Cyber-Physical Systems (with CPS-IoT Week 2023)*. 272–273.

[23] Takahiro A Kato, Shigenobu Kanba, and Alan R Teo. 2019. Hikikomori: multidimensional understanding, assessment, and future international perspectives. *Psychiatry and clinical neurosciences* 73, 8 (2019), 427–440.

[24] Shehroz S Khan, Tiancheng Gu, Lauren Spinelli, and Rosalie H Wang. 2023. Sensor-based assessment of social isolation in community-dwelling older adults: a scoping review. *BioMedical Engineering OnLine* 22, 1 (2023), 1–18.

[25] Juha M Kortelainen, Martin O Mendez, Anna Maria Bianchi, Matteo Matteucci, and Sergio Cerutti. 2010. Sleep staging based on signals acquired through bed

sensor. *IEEE Transactions on Information Technology in Biomedicine* 14, 3 (2010), 776–785.

[26] Scikit learn developers. 2023. ElasticNet. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.ElasticNet.html Online; accessed 3 October 2023.

[27] Scikit learn Developers. 2023. KMeans clustering. https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html Accessed: [Insert Date of Access].

[28] Scikit learn developers. 2023. LinearRegression. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html Online; accessed 3 October 2023.

[29] Scikit learn developers. 2023. RandomForestRegressor. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html Online; accessed 3 October 2023.

[30] Ming-Che Lee, Shu-Yin Chiang, Sheng-Cheng Yeh, and Ting-Feng Wen. 2020. Study on emotion recognition and companion Chatbot using deep neural network. *Multimedia Tools and Applications* 79 (2020), 19629–19657.

[31] Joseph Lee Rodgers and W Alan Nicewander. 1988. Thirteen ways to look at the correlation coefficient. *The American Statistician* 42, 1 (1988), 59–66.

[32] Matthew Li. 2022. *Understanding the Measures of Bluetooth RSSI*. https://www.mokoblue.com/measures-of-bluetooth-rssi/#:~:text=An%20ideal%20RSSI%20value%20of,number%20illustrates%20a%20better%20connection. Accessed: 2023-09-27.

[33] Gosia Lipinska and Kevin GF Thomas. 2019. The interaction of REM fragmentation and night-time arousal modulates sleep-dependent emotional memory consolidation. *Frontiers in psychology* 10 (2019), 1766.

[34] Wei-Yin Loh. 2011. Classification and regression trees. *Wiley interdisciplinary reviews: data mining and knowledge discovery* 1, 1 (2011), 14–23.

[35] Scott M Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. *Advances in neural information processing systems* 30 (2017).

[36] National Academies of Sciences Engineering and Medicine. 2020. *Social isolation and loneliness in older adults: Opportunities for the health care system*. National Academies Press.

[37] Siyun Peng and Adam R Roth. 2022. Social isolation and loneliness before and during the COVID-19 pandemic: A longitudinal study of US adults older than 50. *The Journals of Gerontology: Series B* 77, 7 (2022), e185–e190.

[38] Carla M Perissinotto, Irena Stijacic Cenzer, and Kenneth E Covinsky. 2012. Loneliness in older persons: a predictor of functional decline and death. *Archives of internal medicine* 172, 14 (2012), 1078–1084.

[39] Daniel W Russell. 1996. UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of personality assessment* 66, 1 (1996), 20–40.

[40] Patrick Schober, Christa Boer, and Lothar A Schwarte. 2018. Correlation coefficients: appropriate use and interpretation. *Anesthesia & analgesia* 126, 5 (2018), 1763–1768.

[41] Roger Andre Søraa, Pernille Nyvoll, Gunhild Tøndel, Eduard Fosch-Villaronga, and J Artur Serrano. 2021. The social dimension of domesticating technology: Interactions between older adults, caregivers, and robots in the home. *Technological Forecasting and Social Change* 167 (2021), 120678.

[42] Rebecca C Thurston and Laura D Kubzansky. 2009. Women, loneliness, and incident coronary heart disease. *Psychosomatic medicine* 71, 8 (2009), 836.

[43] Clara I Valero, Saverio Gravina, Alessio Ciuffetti, Marco Di Gioia, Raffaele Gravina, Fernando Boronat, and Carlos E Palau. 2021. EBASI: IoT-based emotion and behaviour recognition system against elderly people social isolation. In *International Symposium on Intelligent and Distributed Computing*. Springer, 3–13.

[44] Matthew P Walker. 2009. The role of sleep in cognition and emotion. *Annals of the New York Academy of Sciences* 1156, 1 (2009), 168–197.

[45] Ding-Qiao Wang, Long-Yu Feng, Jin-Guo Ye, Jin-Gen Zou, and Ying-Feng Zheng. 2023. Accelerating the integration of ChatGPT and other large-scale AI models into biomedical research and healthcare. *MedComm–Future Medicine* 2, 2 (2023), e43.

[46] Yasunori Yamada, Kaoru Shinkawa, Miyuki Nemoto, and Tetsuaki Arai. 2021. Automatic Assessment of Loneliness in Older Adults Using Speech Analysis on Responses to Daily Life Questions. *Frontiers in psychiatry* 12 (2021), 712251.

[47] Ryuji Yamazaki, Shuichi Nishio, Hiroshi Ishiguro, Marco Nørskov, Nobu Ishiguro, and Giuseppe Balistreri. 2014. Acceptability of a teleoperated android by senior citizens in Danish society: A case study on the application of an embodied communication medium to home care. *International Journal of Social Robotics* 6 (2014), 429–442.

[48] Bin Yu, Andrew Steptoe, Kaijun Niu, Po-Wen Ku, and Li-Jung Chen. 2018. Prospective associations of social isolation and loneliness with poor sleep quality in older adults. *Quality of Life Research* 27 (2018), 683–691.