INVITATION TO SUBMIT A RESEARCH PROPOSAL ON AN ASHRAE RESEARCH PROJECT

1811-TRP, Determining Occupancy Patterns in Clusters of Buildings with Data Drawn from Web Based Social Media and Mobile Positions

Attached is a Request-for-Proposal (RFP) for a project dealing with a subject in which you, or your institution have expressed interest. Should you decide not to submit a proposal, please circulate it to any colleague who might have interest in this subject.

Sponsoring Committee: TC 1.5, Emerging Computing Applications
Co-sponsored by: TC 7.10, Occupant Behavior in Building Design and Operation

Budget Range: $220,000 may be more or less as determined by value of proposal and competing proposals.

Scheduled Project Start Date: **April 2026** or later.

All proposals must be received at ASHRAE Headquarters by 8:00 AM, EDT, December 15th, 2025. NO EXCEPTIONS, NO EXTENSIONS. Electronic copies must be sent to rpbids@ashrae.org. Electronic signatures must be scanned and added to the file before submitting. The submission title line should read: 1811-TRP, Determining Occupancy Patterns in Clusters of Buildings with Data Drawn from Web Based Social Media and Mobile Positions, and *"Bidding Institutions Name"* (electronic pdf format, ASHRAE's server will accept up to 10MB)

If you have questions concerning the Project, we suggest you contact one of the individuals listed below:

For Technical Matters
Technical Contact
Bing Dong
Phone:210-458-8189
Email: bidong@syr.edu

For Administrative or Procedural Matters:
Manager of Research & Technical Services (MORTS)
Steve Hammerling
ASHRAE, Inc.
180 Technology Parkway, NW
Peachtree Corners, GA 30092
Phone: 404-636-8400
Fax: 678-539-2111
E-Mail: MORTS@ashrae.net

Contractors who plan to submit a proposal must notify the Manager of Research and Technical Services (MORTS) via email by December 1st. This will ensure that they receive any late or additional information regarding the RFP before the bid due date. Monday, December 1st, 2025 is the deadline for submitting technical inquiries.

All proposals must be submitted electronically. Electronic submissions require a PDF file containing the complete proposal preceded by signed copies of the two forms listed below in the order listed below. **ALL electronic proposals are to be sent to rpbids@ashrae.org.**

All other correspondence must be sent to ddaniel@ashrae.org and shammerling@ashrae.org. In all cases, the proposal must be submitted to ASHRAE by 8:00 AM, EDT, Monday, December 15th 2025.
NO EXCEPTIONS, NO EXTENSIONS.

The following forms (Application for Grant of Funds and the Additional Information form have been combined) must accompany the proposal:

(1) ASHRAE Application for Grant of Funds (electronic signature required) and
(2) Additional Information for Contractors (electronic signature required) ASHRAE Application for Grant of Funds (signed) and

ASHRAE reserves the right to reject any or all bids.

State of the Art (Background)

Over the past few decades, research on occupant behavior has primarily relied on single motion or contact sensors to investigate occupant presence or interactions with building systems. Recently, the vast development in the Internet of Things (IoT) has expanded occupant sensing and data acquisition beyond a single technology. Image-based, threshold and mechanical, motion sensing, radio-based, human-in-the-loop, and consumption sensing are commonly used sensing approaches to collect occupancy data. These technologies have been applied in occupant behavior research, including but not limited to occupant presence, people counting, mobility around the point of interest (POI), human building interactions such as turn on/off lights, thermostats, and window blinds adjustment, energy consumption impacts of miscellaneous loads, and tracking movement. However, the practical challenges of high costs and limited scalability prevent widespread implementation of these sensors in a large number of buildings.

Widely used social media, including X (formerly known as Twitter) and Facebook, make it feasible to mine recorded data to learn building occupancy patterns at a low cost. In other fields, such data has been analyzed to provide new insights in areas including travel recommendation (Majid et al. 2013; Xiang and Gretzel 2010), industrial competitive analysis (He et al. 2013), and activeness identification (Cheong et al. 2011). Majid et al. (2013) collected data from social networks. They found tourist locations from geotag photos. By data analysis and text mining, they established the profile location and built a database. The research defined a matrix to depict the similarities of different users, and then the algorithm could make recommendations when a user came to another city. Chen Y et al. (2018) collected hourly real-time Tencent user density data to analyze the time-spatial distribution of urban park users. The research also introduced POI, accessibility, and surrounding impacts into consideration, using regression analysis to find impact factors. Yang et al. explored human mobility patterns using geotagged social media data. The perspective included activity-time and check-in POI types (Yang et al., 2018). The research analyzed the impact of holidays and gender on behavior patterns. Suaysom introduced a new method called "iterative matrix factorization method" to predict location (Suaysom, 2018). In the research, the additional features include separate clustering and feature prediction approaches. Cheong et al. used the data collected from Facebook and Twitter during the Australian floods to recognize the active users and distinguish the validity of the information (Cheong et al. 2011). The approach included social network analysis and visual network analysis. Xiang et al. used content analysis and multivariate analysis to analyze the factors that influence tourists' choice of destination (Xiang and Gretzel 2010). He et al. took the Pizza industry as a case study to conduct competitive analysis and text mining (He et al. 2013). The preference and habit of pizza were obtained. Pohl et al. proposed batch-based active learning to use social media data for crisis management (Pohl et al., 2018). The algorithm is called batch-based active learning algorithm (OBAL), and it introduced budget into the analysis, including time, cost, and quantity. Wang et al. defined attractiveness as an index to evaluate the forest park design (Wang et al., 2018). The research also compared the data from social media and survey (questionnaire). Kang et al. utilize mobile positioning data from social media platforms to extract typical weekly occupancy profiles of non-residential buildings by cluster analysis (Kang et al. 2021). With the help of POI (point of interest) data and location-based social networks, applications for recommendation are also enriched (Islam et al., 2022). A translation-based knowledge graph-enhanced multi-task learning framework is proposed to improve the unexplored POI recommendation by Hu (Hu B et al. 2022). Table 1 summarizes the state-of-the-art social media data mining research in terms of data and approaches.

In summary, research has been conducted on occupancy mobility, travel patterns, and total annual numbers for buildings using social media data. However, there is a noticeable gap in realistic occupancy schedules for various building types as defined by the ASHRAE Building Energy Quotient. In addition, there is a lack of a systematic method to analyze social media data to derive schedules at different time scales (real-time, daily, weekly, monthly, seasonal, and annually) for both building design and operational purposes.

Table 1 Summary of State-of-the-art

No.	Author	Title	Research	Data	Approach
1	Majid et al. 2013	A context-aware personalized travel recommendation system based on geotagged social media data mining	User preference, tour destination recommendation	Photos and text with geotag	DBSCAN, TF-IDF
2	Chen Y et al. 2018	Emerging social media data on measuring urban park use	Evaluation of park use	Hourly real-time Tencent user density (RTUD) data	Time-spatial distribution analysis; Linear regression
3	Yang C et al. 2018	Exploring human mobility patterns using geo-tagged social media data at the group level	Human mobility patterns	Weibo data with location	Log in analysis, regression analysis
4	Suaysom N 2018	Iterative Matrix Factorization Method for Social Media Data Location Prediction	Location prediction	Tweets data	Iterative Location Oriented Nonnegative Matrix Factorization
5	Cheong F et al. 2011	Social Media Data Mining: A Social Network Analysis Of Tweets During The 2010-2011 Australian Floods	Active user recognition; information validity evaluation	Facebook and Twitter text data	SNA（social network analysis）; visual network analysis
6	Xiang Z et al. 2010	Role of social media in online travel information search	Influence of web on tourist	Query data	Content analysis; Multivariate analysis
7	He W et al. 2013	Social media competitive analysis and text mining: A case study in the pizza industry	Impact of social media on food industry	Facebook and Twitter text data	Text Mining (Based on SPSS and Nvivo 9)
8	Pohl D et al. 2018	Batch-based active learning: Application to social media data for crisis management	Crisis management	Social media data	batch-based active learning algorithm (OBAL)
9	Wang Z et al. 2018	Comparing Social Media Data and Survey Data in Assessing the Attractiveness of Beijing Olympic Forest Park	Evaluation attractiveness of park	Comment data and survey	content data analysis
10	Kang X et al. 2021	Typical weekly occupancy profiles in non-residential buildings based on mobile positioning data	Human mobility patterns	Position data from mobile APP position	Clustering analysis

Table 1 Summary of State-of-the-art

No.	Author	Title	Research	Data	Approach
1	Majid et al. 2013	A context-aware personalized travel recommendation system based on geotagged social media data mining	User preference, tour destination recommendation	Photos and text with geotag	DBSCAN, TF-II
2	Chen Y et al. 2018	Emerging social media data on measuring urban park use	Evaluation of park use	Hourly real-time Tencent user density (RTUD) data	Time-spatial distribution analysis; Linea regression
3	Yang C et al. 2018	Exploring human mobility patterns using geo-tagged social media data at the group level	Human mobility patterns	Weibo data with location	Log in analysis regression analy
4	Suaysom N 2018	Iterative Matrix Factorization Method for Social Media Data Location Prediction	Location prediction	Tweets data	Iterative Locatic Oriented Nonnegative Matrix Factorization
5	Cheong F et al. 2011	Social Media Data Mining: A Social Network Analysis Of Tweets During The 2010-2011 Australian Floods	Active user recognition; information validity evaluation	Facebook and Twitter text data	SNA（social network analysis）; visual networl analysis
6	Xiang Z et al. 2010	Role of social media in online travel information search	Influence of web on tourist	Query data	Content analysi Multivariate analysis
7	He W et al. 2013	Social media competitive analysis and text mining: A case study in the pizza industry	Impact of social media on food industry	Facebook and Twitter text data	Text Mining (Based on SPS and Nvivo 9)
8	Pohl D et al. 2018	Batch-based active learning: Application to social media data for crisis management	Crisis management	Social media data	batch-based acti learning algorith (OBAL)
9	Wang Z et al. 2018	Comparing Social Media Data and Survey Data in Assessing the Attractiveness of Beijing Olympic Forest Park	Evaluation attractiveness of park	Comment data and survey	content data analysis
10	Kang X et al. 2021	Typical weekly occupancy profiles in non-residential buildings based on mobile positioning data	Human mobility patterns	Position data from mobile APP position	Clustering analy
11	Gu J et al. 2018	Extracting typical occupancy data of different buildings from mobile positioning data	Human mobility patterns	Mobile positioning data	k-means clusteri

12	Hu B et al. 2022	TransMKR: Translation-based knowledge graph enhanced multi-task point-of-interest recommendation	Recommending unexplored POIs	Real world check-in datasets	Translation-based knowledge graph enhanced multi-task learning framework (TransMKR)
13	Seo Y et al. 2022	Point of interest recommendations based on the anchoring effect in location-based social network services	Point of interest recommendation	Three datasets from Gowalla (Cho et al., 2011)	A point of interest recommender system based on an anchoring effect
14	Song et al. 2023	Exploring prior knowledge from human mobility patterns for POI recommendation	Point of interest recommendation	Yellow taxi data and social media data	Data structure of graphs
15	Yang et al. 2024	Influence of residential built environment on human mobility in Xining: A mobile phone data perspective	Human mobility patterns	Mobile phone data	Regression analysis on daily movement distance, the radius of gyration (ROG), and the number of stops

Justification and Value to ASHRAE

The proposed research will advance the knowledge and understanding of realistic occupancy presence and patterns in clusters of buildings in a cost-effective way. Such understanding will provide energy modelers with more realistic occupancy schedule inputs, thereby improving urban-scale energy simulation accuracy. This understanding will also improve the accuracy of large-scale load forecasting and building energy simulation. It will also help utility companies develop city planning and manage energy efficiency programs. The outcomes from this research will contribute to the following Chapters in ASHRAE Handbooks:

• Application Handbook: Chapter 61– Smart Buildings Systems

It is aligned with ASHRAE Strategic Plans and Initiatives (2019-2025):

Initiative 1: Resiliency and Decarbonization in Buildings

Initiative 2: Indoor Environmental Quality

Objectives

The overall objectives of this research project are: 1) develop a method to extract social media data with regard to occupancy information; 2) develop algorithms to learn occupancy patterns, specifically presence, in at least 4 types of buildings; 3) establish a typical occupancy profile database for different building types.

Scope:

The scope (technical approach) compromises four tasks in accordance with the four study objectives, the structure figure is as follows:

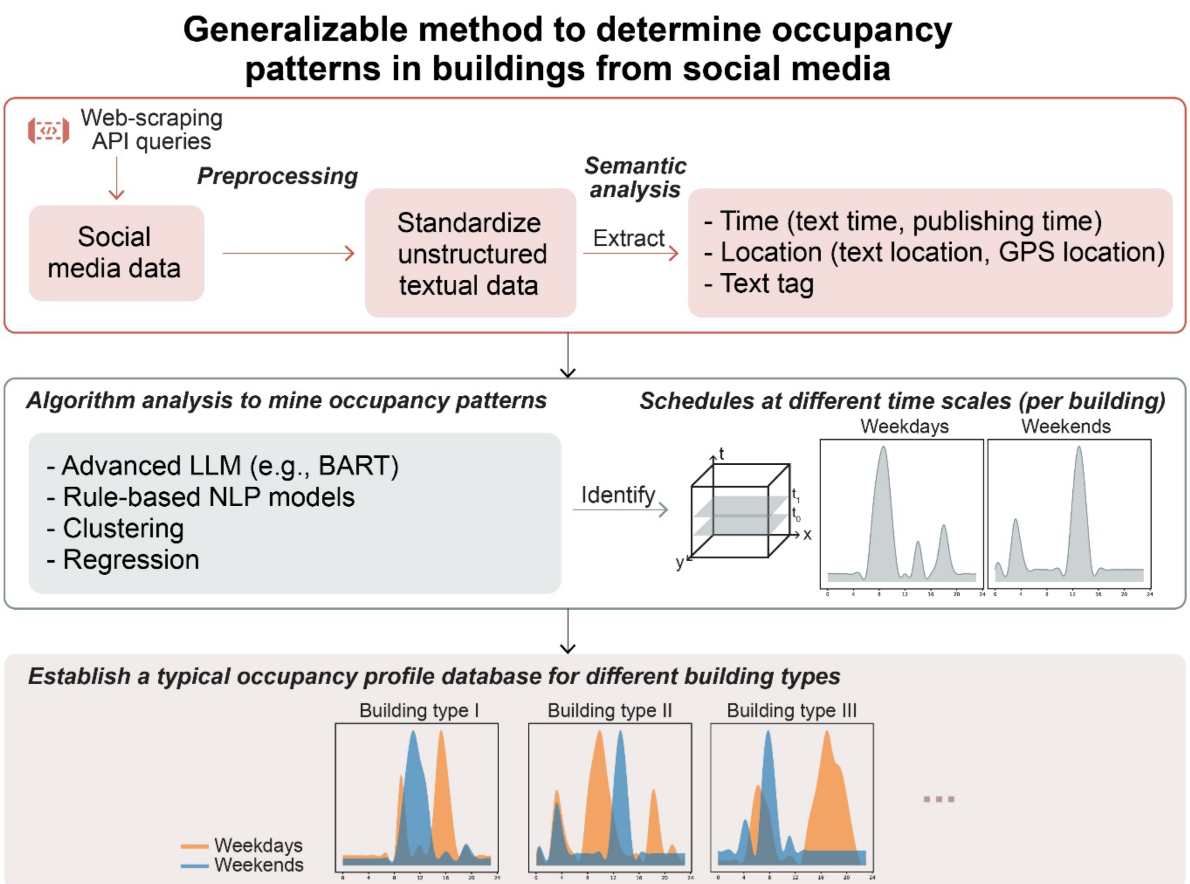

Figure 1 Structure of approach

Task #1: Summarize the evidence that crowdsourced social media data is valuable for urban-scale occupant behavior extraction

With the rapid development of information and communications technology (ICT), urban geolocated data from social media (e.g., Facebook, X, Google Reviews, etc.) promises to expand our understanding of where people are and what they do. However, **to date there is no holistic summary to synthesize the opportunities and evidence that social media data has been used, analyzed,** and reported for the rigorousness and soundness of occupant behavior modeling at urban scale.

Textual data available on social network enriched with multi-layered details, including time (text time or publishing time), location information (text location, GPS location), and the part-of-speech tagging for topic-specific opinion mining, activities, sentiment polarities. There is a need to consolidate the current body of knowledge by collecting all the empirical evidence from eligible studies. To this end, a literature search will be conducted using the Web of Science and Scopus search engines. The keyworks will be used in the literature search including "social media* OR social network* OR social platform*" and "urban* OR city OR cities OR built environments*" and "occupant* OR human behavior OR behavioral*". Specifically, the search formula will be TS (including titles, abstracts, keywords, and extended subject terms) = ("social media* OR social network* OR social platform*") AND (urban* OR city OR cities OR built environments*) AND ("occupant* OR human behavior OR behavioral*"), focusing on studies from the period 2000 to 2024. The retrieved literature will undergo a secondary review, filtering studies that are (1) closely related to occupant behavior topics and (2) clearly utilize social media data in their analysis. The selected empirical studies will be thoroughly examined and summarized under the following categories: (1) Key applications of social media data in supporting occupant behavior modeling, (2) Standardization methods of consolidating and utilizing different social media data sources to enhance occupant behavior modeling, (3) Analytical methods that have been effectively implemented and validated on social media data for behavior classification, clustering, regression, and association analysis, (4) Discussion on opportunities for urban-scale occupant behavior model development.

The preprocessing work for summarizing the evidence, in general, can be concluded as the following three steps:
1) Search the literature and collect relevant evidence to demonstrate the role of social media data in occupant behavior research.
2) Summarize the findings of empirical evidence based on the categories listed above.
3) Provide methodological foundation for the completion of Task 2.

Interim Deliverables: Relevant studies selection and evidence summary that outlines the methods for collecting and preprocessing textual data from social networks, including details on time, location information, and part-of-speech tagging for specific analyses like opinion mining, activities, and sentiment polarities.

PMS Review Points: Critical summary and documentation of the methodology and conclusions derived from empirical evidence.

Task #2 Develop a text mining method to extract occupancy information from social media data

A multi-step approach is required to develop a text mining method for extracting occupancy information from social media data. First, data collection will be conducted either through scraping social media platforms or utilizing APIs to systematically gather posts in the past five years. Following this, the data preprocessing involves a series of steps to clean and normalize the text. This includes removing unnecessary elements such as hashtags, removal of stop words, URLs, and special characters, and standardizing the text through tokenization and potentially lemmatization. Various text mining algorithms (e.g., transformer-based deep learning language models, advanced LLM like BART, rule-based NLP models) will be fine-tuned and implemented to perform semantic analysis on unstructured textual data from social media, extracting and segmenting occupancy information into categories such as text time, publishing time, text location and GPS location. The model's performance will be evaluated using metrics such as precision, recall, and F1 score. Additionally, a qualitative analysis will be conducted to confirm that the segmented occupant behavior information aligns with the expected outcomes and maintains a high level of relevance and contextuality. Since the text data are not temporally continuous, the location data of all messages cannot represent a real total number of people in a specific building or in a cluster of buildings. Therefore, to improve accuracy, regression and correlation analysis of this kind of data must be performed. Specifically, changes in a single user's location can be tracked without disclosing any identifying information. This allows for the analysis of time tags and the locations by means of regression analysis.

The data from the location service can be considered temporally and spatially continuous. Hence, by data cleaning, the total number of occupants in a building or in a cluster of buildings can be estimated. Typical clustering methods such as Support Vector Machines (SVM), t-distributed stochastic neighbor embedding, and time-series prediction methods such as Artificial Neural Network (ANN) shall be customized and developed for this project. Through that analysis, we specifically can derive dynamic occupancy profiles that represent the varying occupancy trends across different timescales: daily, weekly, monthly, and seasonal. Currently, the ASHRAE handbook uses a fixed schedule for HVAC design and sizing in commercial buildings throughout the year. This static approach can lead to discrepancies between the design values and the actual operational values, influencing peak load, the timing of peak load occurrences, and total load. By leveraging social media data and mobile positioning data from the past five years, our methodology allows us to identify large-scale patterns and trends that inform more accurate, predictive modeling of occupancy-related demands on more energy-efficient scheduling of HVAC systems.

In addition to occupancy prediction, the occupancy amount presents periodicity and regularity, so time-series prediction methods can be used to analyze the routine with occupancy and predict the occupancy amount in the next step. The potential algorithms include the ARIMA (Autoregressive Integrated Moving Average) model and the Holt-Winters model. The occupancy prediction can be used for different objectives. If the resolution is around minutes, one, or two hours, then the prediction may help with building operation and control. If the resolution is around one day, then the prediction may be used as cooling load prediction, to decide the amount of next day cooling load.

It is worth noting that the social media data referenced in this project includes a variety of platforms, including X and Facebook, alongside mobile phone networks. This diverse range of sources captures a wide spectrum of user demographics, providing a robust method for obtaining urban-scale occupancy profiles. Such granularity would be challenging to achieve using occupancy sensors and surveys alone. In this study, we specifically focus on buildings such as offices, museums, shopping malls, and hotels, which demonstrate large uncertainties in their occupancy patterns. In these venues, age groups that typically do not use phones often visit alongside those that do, thereby enhancing the reliability and coverage of the social media data.

Interim Deliverables: A literature review on different data mining methods on occupancy data processing. Confirm three methods for occupancy pattern calculation.

PMS Review Points: Full set of analysis methods and relative analysis method program.

<u>Task #3 Determine a typical occupancy profile database for different building types</u>

This task focuses on creating a typical occupancy profile database for various building types, using the information extracted in Task #2. This database aims to standardize occupancy profiles across different categories of buildings, such as offices, museums, shopping malls, and hotels.

The creation of this database begins with the aggregation of extracted occupancy data from at least 100 buildings across similar types of buildings. Statistical and machine learning methods, such as Fourier analysis and functional regression, are employed to establish typical occupancy profiles. These profiles reflect the representative occupancy trends across various timescales, such as daily, weekly, monthly, or seasonal variations. For each building category, the database includes key metrics such as average occupancy, peak occupancy times, and minimum occupancy periods. This information is crucial for stakeholders in building management, urban planning, and emergency services, providing them with a reliable benchmark. Additionally, the database is designed to be dynamic, allowing updates as new data becomes available. This ensures that the occupancy profiles remain relevant and reflective of current usage patterns.

Interim Deliverables: 1) Mid-term Report: This report will detail the aggregation of extracted occupancy data from at least 100 buildings. It should include a summary of the data sources, the types of buildings analyzed, and any initial observations or trends such as preliminary occupancy profiles. The report will include the typical occupancy profiles for each building category, including key metrics such as average occupancy, peak occupancy times, and minimum occupancy periods; 2) Amendment Method on Raw Data: This will describe techniques for data cleaning, handling unstructured social media data, and ensuring data consistency. The goal is to maintain the integrity and reliability of the raw data before applying statistical analysis to derive the typical occupancy profiles.

PMS Review Points: Full edition of validation and evaluation report.

Task #4 Validating social media-driven occupancy profile in a real building

This task focuses on validating social media-driven occupancy profiles. The contractor will select a type of building such as a research office building or a library, and collect its social media data for at least six months, alongside ground truth building occupancy data (e.g., using people counting sensors). Specifically, the contractor shall encourage users to check in or use location tags on platforms like Facebook or X when they enter or spend time in the building. The contractor shall encourage them to post content related to their visit (e.g., duration and purpose) or specific events hosted there by using predetermined hashtags. The contractor will also monitor mentions and discussions related to events or facilities within the building, such as special exhibits in the library or guest lectures in research centers.

To validate the occupancy data derived from social media against the "ground truth" data collected via traditional methods, the contractor will employ the following methods:
(1) Measure the Mean Absolute Error (MAE) between the occupancy estimates from social media and sensor data to provide a direct average measure of how much the social media data deviates from the sensor data;
(2) Categorize occupancy into categories (e.g., low, medium, high) and use classification metrics like the F1 Score to evaluate the accuracy of these categorizations against sensor data;
(3) Conduct cross-correlation analysis to assess the relationship between social media indicators and actual occupancy levels over time. We will use cross-correlation functions to identify the time lags at which social media data most closely correlates with sensor-based occupancy data, helping to validate the effectiveness of social media as a proxy indicator.

Interim Deliverables: 1) Mid-term Report: This report will provide a detailed account of the data collected from social media and sensors, including the volume of check-ins, location tags, and mentions, documentation of statistical methods and data collection methods, initial assessment of social media-derived data and sensor data, and any preliminary findings derived from the methodology validation procedures. 2) Data Quality and Integrity Review: This document will evaluate the quality and reliability of the data collected, identifying any inconsistencies or gaps in both social media and sensor data. It will also include strategies for data cleaning and preprocessing to ensure robust analysis. Recommendations for improving data collection and validation methods based on these findings will be outlined to enhance the accuracy of future measurements and predictions.

PMS Review Points: Full edition of validation and evaluation report.

Deliverables:

Progress, Financial and Final Reports, Technical Paper(s), and Data shall constitute the deliverables ("Deliverables") under this Agreement and shall be provided as follows:

a. Progress and Financial Reports

Progress and Financial Reports, in a form approved by the Society, shall be made to the Society through its Manager of Research and Technical Services at quarterly intervals; specifically on or before each January 1, April 1, June 10, and October 1 of the contract period.

The following deliverables shall be provided to the Project Monitoring Subcommittee (PMS) as described in the Scope/Technical Approach section above, as they are available:

Furthermore, the Institution's Principal Investigator, subject to the Society's approval, shall, during the period of performance and after the Final Report has been submitted, report in person to the sponsoring Technical Committee/Task Group (TC/TG) at the annual and winter meetings, and be available to answer such questions regarding the research as may arise.

b. Final Report

A written report, design guide, or manual, (collectively, "Final Report"), in a form approved by the Society, shall be prepared by the Institution and submitted to the Society's Manager of Research and Technical Services by the end of the Agreement term, containing complete details of all research carried out under this Agreement, including a summary of the control strategy and savings guidelines. Unless otherwise specified, the final draft report shall be furnished, electronically for review by the Society's Project Monitoring Subcommittee (PMS).

Tabulated values for all measurements shall be provided as an appendix to the final report (for measurements which are adjusted by correction factors, also tabulate the corrected results and clearly show the method used for correction).

Following approval by the PMS and the TC/TG, in their sole discretion, final copies of the Final Report will be furnished by the Institution as follows:

-An executive summary in a form suitable for wide distribution to the industry and to the public.
-Two copies; one in PDF format and one in Microsoft Word.

c. ASHRAE Handbooks

Updates to appropriate sections of ASHRAE Handbooks shall be proposed to the relevant TC's Handbook Subcommittee Chairs.

d. *Science & Technology for the Built Environment*

One or more papers shall be submitted first to the ASHRAE Manager of Research and Technical Services (MORTS) and then to the "ASHRAE Manuscript Central" website-based manuscript review system in a form and containing such information as designated by the Society suitable for publication. Papers specified as deliverables should be submitted to Research Papers for HVAC&R Research for ASHRAE Transactions. Research papers contain generalized results of long-term archival value, whereas technical papers are appropriate for applied research of shorter-term value, ASHRAE Conference papers are not acceptable as deliverables from ASHRAE research projects. The paper(s) shall conform to the instructions posted in "Manuscript Central" for HVAC&R Research papers. The paper title shall contain the research project number (1811-RP) at the end of the title in parentheses, e.g., (1811-RP).

All papers or articles prepared in connection with an ASHRAE research project, which are being submitted for inclusion in any ASHRAE publication, shall be submitted through the Manager of Research and Technical Services first and not to the publication's editor or Program Committee.

e. Data

Data is defined in General Condition VI, "DATA"

f. Project Synopsis

A written synopsis totaling approximately 100 words in length and written for a broad technical audience, which documents 1. Main findings of research project, 2. Why findings are significant, and 3. How the findings benefit ASHRAE membership and/or society in general shall be submitted to the Manager of Research and Technical Services by the end of the Agreement term for publication in ASHRAE Insights

The Society may request the Institution submit a technical article suitable for publication in the Society's ASHRAE JOURNAL. This is considered a voluntary submission and not a Deliverable. Technical articles shall be prepared using dual units; e.g., rational inch-pound with equivalent SI units shown parenthetically. SI usage shall be in accordance with IEEE/ASTM Standard SI-10.

Level of Effort
The project is expected to take 24 months with an estimated budget of $220,000. It will require at least 3 person-months for the principal investigator and 24 person-months for a research assistant. Estimated effort breakdown by task:

Other Information to Bidders (Optional):

1. Background Literature Review: 5%
2. Method Development: 20%
3. Determine a typical occupancy profile database: 35%
4. Field validation: 35%
5. Reporting: 5%

Proposal Evaluation Criteria:

No.	Proposal Review Criterion	Weighting Factor
1	Contractor's understanding of Work Statement as revealed in the proposal.	20%
2	Qualification of personnel for this project.	30%
3	Quality of methodology proposed for conducting research.	30%
4	Organization of the proposal	10%
5	Involvement of students	10%

Project Milestones:

No.	Major Project Completion Milestone	Deadline Month
1	**Task 1:** Screen and summarize related literature evidence	2 mo. from start
2	**Task 2:** Develop a data collection method for social media/mobile position data	4 mo. from start
3	**Task 3:** Process and normalize the collected social media/mobile position data	8 mo. from start
4	**Task 4:** Implement, fine-tune, and examine various text mining algorithms for occupancy pattern extraction	10 mo. from start
5	**Task 5:** Identify a typical occupancy profile for different building types	16 mo. from start
6	**Task 6:** Establish an occupancy profile database	18 mo. from start
7	**Task 7:** validate occupancy profile using one building data	24 mo. from start
8	**Final Deliverables:** Final task reports and data sets, final summary report, technical paper, Project synopsis	24 mo. from start

References

1. Chang, Wen-kuei, and Tianzhen Hong. 2013. Statistical analysis and modeling of occupancy patterns in open-plan offices using measured lighting-switch data. Building Simulation, 6(1): 23-32.
2. CHEN, Y. Y., LIU, X. P., GAO, W. X., WANG, R. Y., LI, Y. & TU, W. 2018. Emerging social media data on measuring urban park use. Urban Forestry & Urban Greening, 31, 130-141.
3. Cheong, F, and C Cheong. 2011. Social media data mining: A social network analysis of tweets during the Australian 2010-2011 floods. In 15th Pacific Asia Conference on Information Systems (PACIS), 1-16. Queensland University of Technology.
4. D'Oca, Simona, and Tianzhen Hong. 2014. A data-mining approach to discover patterns of window opening and closing behavior in offices. Building and Environment, 82: 726-739.

5. Eilers, Mark, John Reed, and TecMRKT Works. 1996. Behavioral aspects of lighting and occupancy sensors in private offices: a case study of a university office building. ACEEE 1996 Summer Study on Energy Efficiency in Buildings.
6. Erickson, Varick L, Miguel A Carreira-Perpinan, and Alberto E Cerpa. 2011. OBSERVE: Occupancy-based system for efficient reduction of HVAC energy. In Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on, 258-269. IEEE.
7. Fan, C., Yan, D., Xiao, F., Li, A., An, J., & Kang, X. (2021). Advanced data analytics for enhancing building performances: From data-driven to big data-driven approaches. Building Simulation, 14(1), 3-24.
8. Feng, Xiaohang, Da Yan, Chuang Wang, and Hongsan Sun. 2016. A preliminary research on the derivation of typical occupant behavior based on large-scale questionnaire surveys. Energy and Buildings, 117: 332-340.
9. Gu, J., Xu, P., Pang, Z., Chen, Y., Ji, Y., & Chen, Z. (2018). Extracting typical occupancy data of different buildings from mobile positioning data. Energy and Buildings, 180, 135-145.
10. Guo, X, DK Tiller, GP Henze, and CE Waters. 2010. The performance of occupancy-based lighting control systems: A review. Lighting Research and Technology, 42(4): 415-431.
11. He, Wu, Shenghua Zha, and Ling Li. 2013. Social media competitive analysis and text mining: A case study in the pizza industry. International Journal of Information Management, 33(3): 464-472.
12. Hu, B., Ye, Y., Zhong, Y., Pan, J., & Hu, M. (2022). TransMKR: Translation-based knowledge graph enhanced multi-task point-of-interest recommendation. Neurocomputing, 474, 107-114.
13. Islam, M. A., Mohammad, M. M., Sarathi Das, S. S., & Ali, M. E. (2022). A survey on deep learning-based Point-of-Interest (POI) recommendations. Neurocomputing, 472, 306-325.
14. Kang, X., Yan, D., An, J., Jin, Y., & Sun, H. (2021). Typical weekly occupancy profiles in non-residential buildings based on mobile positioning data. Energy and Buildings, 250, 111264.
15. Leephakpreeda, T, R Thitipatanapong, T Grittiyachot, and V Yungchareon. 2001. Occupancy-based control of indoor air ventilation: A theoretical and experimental study. Science Asia, 27(4): 279{84.
16. Lo, L James, and Atila Novoselac. 2010. Localized air-conditioning with occupancy control in an open office" Energy and Buildings, 42(7): 1120-1128.
17. Majid, Abdul, Ling Chen, Gencai Chen, Hamid Turab Mirza, Ibrar Hussain, and John Woodward. 2013. A context-aware personalized travel recommendation system based on geotagged social media data mining. International Journal of Geographical Information Science, 27(4): 662-684.
18. Page, Jessen, Darren Robinson, Nicolas Morel, and J-L Scartezzini. 2008. A generalised stochastic model for the simulation of occupant presence. Energy and buildings, 40(2): 83-98.
19. POHL, D., BOUCHACHIA, A. & HELLWAGNER, H. 2018. Batch-based active learning: Application to social media data for crisis management. Expert Systems with Applications, 93, 232-244.
20. Richardson, Ian, Murray Thomson, and David Infield. 2008. A high-resolution domestic building occupancy model for energy demand simulations. Energy and buildings, 40(8): 1560-1566.
21. Seo, Y.-D., & Cho, Y.-S. (2021). Point of interest recommendations based on the anchoring effect in location-based social network services. Expert Systems with Applications, 164, 114018.
22. SUAYSOM, N. 2018. Iterative Matrix Factorization Method for Social Media Data Location Prediction.
23. Wang, Chuang, Da Yan, and Yi Jiang. 2011. A novel approach for building occupancy simulation. Building Simulation, 4(2): 149-167.
24. Wang, Danni, Clifford C Federspiel, and Francis Rubinstein. 2005. Modeling occupancy in single person offices." Energy and buildings, 37(2): 121-126.
25. WANG, Z., JIN, Y., LIU, Y., LI, D. & ZHANG, B. 2018. Comparing Social Media Data and Survey Data in Assessing the Attractiveness of Beijing Olympic Forest Park. Sustainability, 10, 382.
26. Xiang, Zheng, and Ulrike Gretzel. 2010. Role of social media in online travel information search. Tourism management, 31(2): 179-188.
27. Yamaguchi, Yohei, Takuya Fujimoto, and Yoshiyuki Shimoda. 2011. Occupant behavior model for households to estimate high-temporal resolution residential electricity demand profile." Proceedings of BS2011, 1548-1555.
28. YANG, C., XIAO, M., DING, X., TIAN, W., ZHAI, Y., CHEN, J., LIU, L. & YE, X. 2018. Exploring human mobility patterns using geo-tagged social media data at the group level. Journal of Spatial Science, 1-18.
29. Yu, Tina. 2010. Modeling occupancy behavior for energy efficiency and occupants comfort management in intelligent buildings. In Machine Learning and Applications (ICMLA), 2010 Ninth International Conference on, 726-731. IEEE.

30. Yu, Zhun, Benjamin CM Fung, Fariborz Haghighat, Hiroshi Yoshino, and Edward Morofsky. 2011. A
31. **systematic procedure to study the influence of occupant behavior on building energy consumption. Energy and Buildings, 43(6): 1409-1417.**