CausalCFF: Causal Analysis between User Stress Level and Contextually Filtered Features Extracted from Mobile Sensor Data

Panyu Zhang
Graduate School of Data Science
KAIST
Daejeon, Republic of Korea
panyu@kaist.ac.kr

Gyuwon Jung
School of Computing
KAIST
Daejeon, Republic of Korea
gwjung@kaist.ac.kr

Uzair Ahmed
School of Computing
KAIST
Daejeon, Republic of Korea
uziahmd@kaist.ac.kr

Uichin Lee
School of Computing
KAIST
Daejeon, Republic of Korea
uclee@kaist.edu

Abstract

Nowadays, it's possible to deliver interventions through mobile technologies to improve users' mental and physical health. Causal analysis may help researchers identify the potential causes of the health issues and design proper interventions. However, in previous studies, causal analysis is mainly conducted between single sensor data features such as walking activity duration and perceived stress. There is a lack of research into causal analysis between more complicated behavior features which could be derived from multiple sensor features and target well-being labels. To address this gap, we propose **CausalCFF**, a framework that investigates causal relationships between contextually filtered behavioral features (e.g., walking duration at workplace locations) and well-being outcomes (e.g., stress). Our analysis identifies frequent workplace visits during periods of reduced home time as the most salient cause for elevated stress levels, highlighting the framework's ability to target context-specific behavioral biomarkers for human well-being. The code is also made available[1].

CCS Concepts

• **Human-centered computing** → **Ubiquitous and mobile computing**; • **Applied computing** → *Health informatics*; • **Computing methodologies** → **Causal reasoning and diagnostics**.

Keywords

Causal Analysis, Contextually Filtered Features, Human Behavior, Mobile Sensor Data, Stress

ACM Reference Format:
Panyu Zhang, Gyuwon Jung, Uzair Ahmed, and Uichin Lee. 2025. CausalCFF: Causal Analysis between User Stress Level and Contextually Filtered Features Extracted from Mobile Sensor Data. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3706599.3719776

[1]https://github.com/Kaist-ICLab/CausalCFF

CHI EA '25, Yokohama, Japan

ACM ISBN 979-8-4007-1395-8/25/04
https://doi.org/10.1145/3706599.3719776

1 Introduction and Background

Mobile technology has revolutionized personal health management, enabling data-driven interventions that support users in understanding and improving their physical and mental well-being [5, 10, 13, 18, 23]. Central to these interventions is the ability to uncover causal relationships between users' behaviors, contexts, and health outcomes, which is critical for designing effective health interventions.

Previous studies have highlighted the importance of examining these relationships to better support users in self-reflection and behavior change. For instance, Mehrotra et al. [16] explored how mobile phone interactions influence emotional states by conducting correlational and causal analyses. Similarly, Kim et al. [11] leveraged correlational insights to help college students reflect on their mental health, providing feedback such as "*When your stress level is high, you stay home all day.*" Jung et al. [8] extended this approach by applying causal inference methods to observational mobile data, enabling users to rigorously investigate which contexts are causally linked to their stress levels. For example, their approach identified whether the 'class-taking' activity causes increased stress while controlling for external confounding factors such as locations, social settings, and time. Building on this foundation, our study focuses on investigating the factors that cause stress—one of the most prevalent mental health challenges in modern society—using mobile data and a novel causal analysis approach.

In the Ubiquitous Computing and Human-Computer Interaction domains, researchers have explored various causal analysis methods to understand how observational mobile sensor data can reveal insights into human well-being. Tsapeli et al. [25] proposed a quasi-experimental approach using *matching* techniques to examine causal relationships between time spent in different locations and students' stress levels. Jung et al.[9] further summarized a matching-based causal analysis pipeline for exploring causal relationships in human behaviors and contexts.

Beyond matching-based approaches, Berkel et al. [26] introduced 'Convergent Cross Mapping (CCM),' a method first developed in environmental studies [22], to analyze causal dynamics in human behavior. CCM is grounded in Taken's theorem [24], which posits that in dynamic systems, if variable X causes variable Y, X's information is included in Y, allowing X to be reconstructed using Y. Unlike matching-based methods, CCM does not focus on estimating the average treatment effect (ATE) but instead reconstructs causal relationships by leveraging embedded information in dynamic systems [7]. Sarsenbayeva et al. [20] applied CCM to examine the causal relationships between smartphone usage and emotions, highlighting its potential to uncover complex, non-linear relationships in human behavior.

Despite these advances, analyzing causal relationships remains challenging due to the complexity of human behavior and context. Most previous studies have focused on single behavioral or contextual features (e.g., activities or locations) to assess their impact on mental or physical health. However, such analyses may overlook the fact that real-world situations are influenced by combinations of multiple features. For instance, phone usage alone may not generally increase or decrease stress, but the same activity in certain contexts (e.g., office) might have a significant effect.

To address these limitations, we propose combining association rule mining with the causal analysis method. Association rule mining can analyze multiple raw sensor data features to extract complex behavior patterns from mobile sensor data [27]. Behavior rules are expressed in the form of consequent | antecedent, where the 'antecedent' represents the condition or context, and the 'consequent' indicates the resulting behavior. Using these rules, 'contextually filtered features' are defined as the mean and standard deviation of the consequent feature given the antecedent. This approach has demonstrated effectiveness in predicting outcomes such as depression and user receptivity [1, 27]. For instance, if the rule "walking activity duration = high (consequent) | location = office (antecedent)" is identified as a frequent pattern with high lift, the respective contextually filtered features can be derived by calculating the mean duration of walking activity when the location is the office.

While sequential rule mining [14] could also be applied to extract human behavior features by considering temporal order, this paper focuses on contextually filtered features for simplicity. This approach serves as an initial step toward conducting causal analysis on complex behavior patterns derived from single sensor data features.

2 Method

2.1 Dataset

We utilize an open dataset from a recent study on understanding causal relationships between contextual factors and perceived stress [8]. This dataset comprises smartphone data, including GPS, physical activity (Activity Recognition APIs), app usage records (Usage Event APIs) and self-reported perceived stress levels on a 5-point Likert scale. Stress labels are gathered using the Experience Sampling Method at intervals of at least 30 minutes. Overall, data are collected from 24 university students over six weeks. We choose this dataset because it includes key features of interest that may be associated with stress. Further details of the dataset can be found at: https://github.com/Kaist-ICLab/DeepStress_Dataset/

In this pilot study, we utilize location, physical activity, and app usage data for behavior feature extraction for all users. The stress label distribution of this dataset is shown in the following Figure 1.

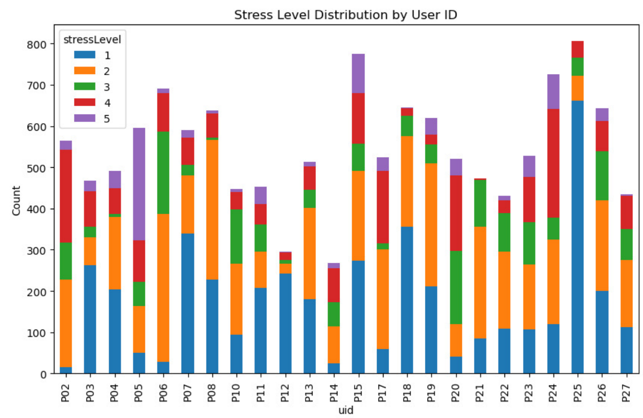

Figure 1: Stress Label Distribution for the DeepStress Dataset

2.2 Data Analysis Pipeline for CausalCFF Framework

2.2.1 Data preprocessing & feature extraction. To ensure the interpretability of the contextually filtered extracted features, we consider only interpretable single sensor data features as candidates. For example, instead of adopting all the possible numeric features of social app usage durations within a time window, we might simply retain their total sum, thereby discarding the skewness of the distribution of app usage durations.

As for location data, we calculate the haversine distance between consecutive GPS recordings and cluster the GPS data for each user using POI clustering [17]. The clusters are labeled using semantic labels (home, work, labels returned by Google Maps API, and none) [25, 28]. Home and work clusters are first labeled mainly considering the time spent at those locations. Rest of the clusters are labeled using Google Maps places API. As for app usage data, we recategorize apps into predefined categories and calculate app usage duration for each category [21]. As for physical activity data, duration for each physical activity type is calculated based on the physical activity transition events. Specifically, the durations are calculated based on transition events such as enter_walking and enter_sitting, and their timestamps accordingly. The detailed feature design can be found in Table 1.

As for missing data handling, missing numeric data are imputed using the mean value, while missing categorical data are filled using forward filling.

2.2.2 Feature Preparation. In order to analyze causal relationships between contextually filtered features and stress labels, there should be more than one context for each stress label so that contextually filtered features can be derived. Therefore, we adopt the approach proposed by Alikhanov et al. [1], using a full time window of 160

Table 1: Feature Design

Data Types	Features	Semantic Meaning
App Usage	'APP_DUR_TYPE#SUM', 'APP_CAT#SUP:TYPE' ('TYPE' denotes app types)	Duration and support/count of app usage for each type. The app usage types include social, system, work, entertainment, information, and health.
Physical Activity	'ACT_DUR_TYPE#SUM' ('TYPE' denotes physical activity types)	Duration of physical activity for each type. The physical activity types include walking, still, in vehicle, on bicycle, and running.
Location	'LOC_DUR_TYPE#SUM', 'LOC_LABEL#SUP:TYPE' ('TYPE' denotes location label types)	Duration and support/count of staying in different types of locations. The location label types include social, none (on the way between locations), work, eating, home, and gym.

minutes and dividing it into 8 sub-time windows to extract features. This ensures that each sub-time window contains sufficient sensor data and that the full time window captures enough contextual variation for each stress label. As shown in the following Figure 2, before each stress label timestamp, we divide 160 minutes full time windows into 8 sub time windows so that there are enough variations of different contexts for the stress label.

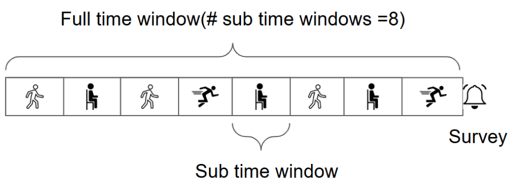

Figure 2: Time Window Setting

Before association rule mining, the features are categorized into three levels including low (l), medium (m), and high (h) based on quantile thresholds 33% and 66%. A transaction list is then constructed using the categorized features.

2.2.3 Association Rule Mining and Contextually Filtered Feature Extraction. The next step involves applying association rule mining to identify the most frequent patterns of routine behavior. To ensure the derived behavior features are generalizable across users, we utilize the entire dataset encompassing all users for behavior feature extraction. This approach is critical because relying on individual user data for behavior feature extraction could result in personalized behavior features specific to each user. Such personalization would hinder our ability to establish a generalized causal relationship between behavior features and stress labels. By using collective data, we aim to uncover broader patterns representative of the population, thereby enabling more robust and generalizable insights.

To limit the number of possible rules, we apply the following criteria: we retain a rule only if its support is above 0.3, its confidence is above 0.8, and the length of both the antecedent and the consequent is fewer than 5 items. Support above 0.3 ensures the rule is applicable to more than 30% of the data, making it statistically significant, while confidence above 0.8 ensures the rule is accurate more than 80% of the time when the antecedent occurs, providing reliability. Our intial threshold selection is guided by Alikhanov et al.'s work [1]. Besides, we experiment with various support thresholds [0.1, 0.2, 0.3, 0.4, 0.5] and confidence thresholds [0.2, 0.4, 0.6, 0.8]. Ultimately, the thresholds are finalized to achieve a balance between rule frequency, relevance, and the total number of extracted rules. For better interpretability, we further filter out rules if their consequent length exceeds 1. For example, we retain 'walking activity duration when location is office,' but exclude 'walking activity duration and entertainment app usage duration when location is office.' The final top features are selected based on their lift values[1]. Lift helps identify significant behavioral associations in human routines by comparing observed and expected co-occurrence frequencies, filtering out random coincidences.

Based on the top rules selected, contextually filtered features can be extracted as shown in the following Figure 3 given the rule heart rate | sitting activity. As shown in the example, among the eight sub time windows, we focus exclusively on those where the context is sitting. For these specific windows, we calculate the mean and standard deviation of the heart rates.

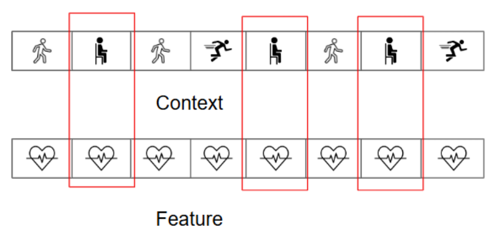

Figure 3: Extraction of Contextually Filtered Features

2.2.4 Causal Analysis. For the causal analysis, we adopt a user-specific approach due to the potential personalization of causal relationships across users. Specifically, we conduct causal analysis individually for each user using CCM, implemented via the causal-ccm package [3]. The causal strength is quantified by computing the correlation between the reconstructed X (derived from Y) and the actual X. This approach is based on the assumption that X causally influences Y, as supported by Takens' theorem [24]. Only correlations that were statistically significant ($p < 0.05$) are included in the overall causality correlation calculation. For each user, the causality correlation is computed individually, and the results are aggregated by calculating the mean causality correlation across all users. This method ensures a robust and generalized understanding of the causal relationships while effectively accounting for individual variability.

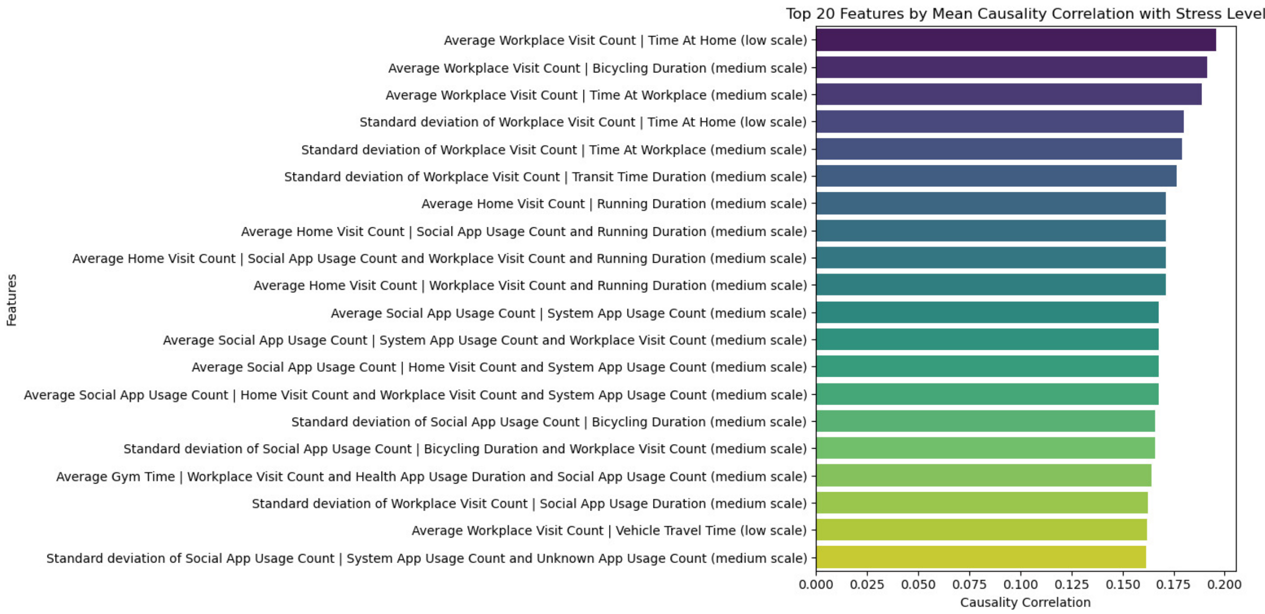

Figure 4: Causal Analysis Results

3 Results

The top 20 contextually filtered features that cause stress for all the users can be found in Figure 4. These features were named before the '|' is the feature while after it is the context condition.

As shown in Figure 4, regardless of the context, the top 20 contextually filtered features that cause stress predominantly include workplace visit count, home visit count, and social app usage count. In particular, the top contextually filtered feature associated with stress is *Average Workplace Visit Count | Time At Home (low scale)*. This finding aligns well with our earlier assumption that human behavior is more complex than what can be captured by a single feature. Even when derived from the same sensor (e.g., GPS data), contextually filtered feature extraction allows us to generate more nuanced representations of user behavior. As for human behavior understanding, this feature suggests that long work hours, combined with limited time at home, may lead to stress.

In the top 20 contextually filtered features, we also observe several multi-modal ones, incorporating data from multiple sensors. For instance, the eighth highest contextually filtered feature, *Average Home Visit Count | Social App Usage Count and Running Duration (medium scale)*, suggests that the frequency of home visits—when coupled with moderate running duration and moderate social app usage—may be associated with higher stress for certain users. These findings align with our research objective of identifying high-level context-specific factors influencing users' stress levels. Such insights can empower researchers in the health behavior change field to design more effective context-aware interventions. By understanding these contextual triggers, we can tailor interventions to specific situations, enabling more personalized and impactful support for users.

Furthermore, the causal relationships demonstrate significant personalization across users, as depicted in Figure 5. Specifically, for the top feature—workplace visit count given the low time spent at home—only 9 out of 24 users exhibit a significant causal relationship between this feature and stress levels. This variability highlights the critical need to account for individual differences when analyzing causal relationships in behavioral data.

4 Discussion

In this study, we proposed **CausalCFF**, a novel approach to uncover causal relationships from observational mobile sensor data by integrating association rule mining and causal analysis. This approach effectively identified stress-inducing factors in the form of contextually filtered features, enabling the pinpointing of direct contributors to stress in scenarios where multiple contextual factors coexist and interact. By eliminating the need to manually test a large number of contextual combinations, this method reduces the cognitive burden associated with identifying stressors. Given the repetitive nature of daily routines and recurring contexts, the contextually filtered features identified through this approach can be analyzed and delivered proactively. Providing these insights in advance can enable users to better manage their stress by avoiding specific contexts or planning appropriate coping strategies. Below, we discuss additional considerations for applying our approach in real-world settings.

4.1 Unveiling Insights into Contextual Features and Stress

As shown in Figure 4, most of the top-ranked contextually filtered features are related to workplace/home visit frequencies and social app usage durations within specific contexts. In particular, the highest-ranked feature suggests that frequent workplace visits, combined with limited time at home, may lead to elevated stress

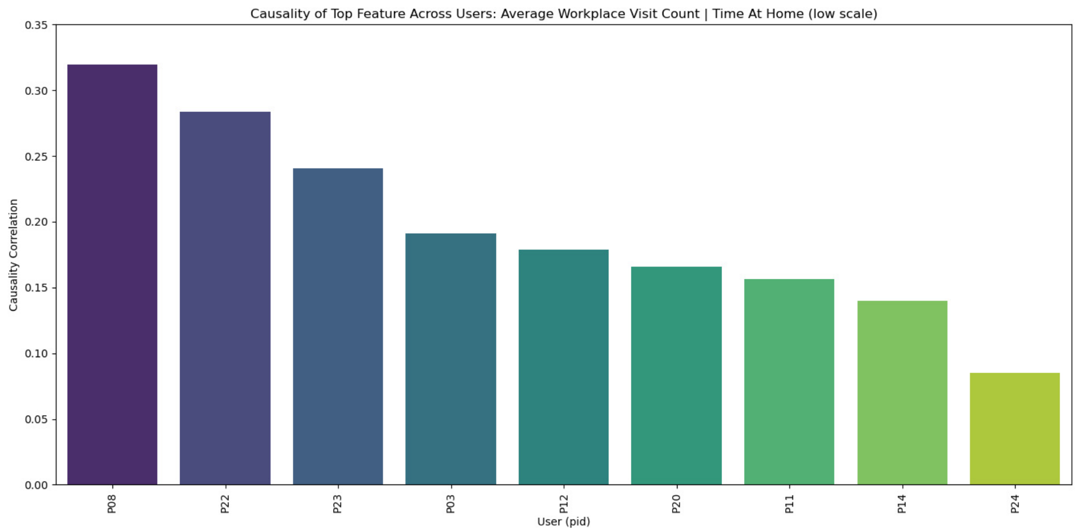

Figure 5: Causality Correlation for Top Feature Across Users

levels. While intuitive, this finding should be interpreted with consideration of individual and contextual factors.

For instance, frequent workplace visits may increase stress for some due to environmental demands, while others might find such visits engaging or routine. Similarly, limited time at home might not inherently lead to stress if the quality of time spent at home (e.g., restful activities or family interactions) provides sufficient recovery. These variations highlight the importance of understanding individual differences and uncovering more specific, hidden contextual factors that may influence these relationships.

The role of social app usage also merits closer investigation, as its effects on stress may vary depending on the app's purpose, intensity of interactions, and the context of its use. Further research should employ mixed methods, such as longitudinal tracking and qualitative studies, to explore how these contextual dynamics influence stress over time.

4.2 Designing Personalized Interventions with CausalCFF

In the domain of health behavior change, designing effective interventions requires balancing theoretical foundations with empirical evidence. Traditional approaches grounded in domain knowledge, such as the *Transtheoretical Model* and *Social Cognitive Theory* [2, 19], provide valuable insights into behavior change processes but may lack the flexibility to address the complexities of individual contexts. Conversely, data-driven approaches leverage patterns from large-scale datasets, uncovering actionable insights and causal relationships that can guide personalized interventions.

Our proposed **CausalCFF** method builds on this data-driven paradigm by uncovering causal relationships specific to user contexts, enabling interventions that dynamically adapt to shifting behavior patterns and contextual changes. This adaptability is crucial in addressing the multifaceted nature of human behavior, where stressors often vary with contextual factors [12]. By integrating contextual causality into intervention design, **CausalCFF** offers a systematic way to tailor strategies to users' immediate needs while also accounting for longer-term behavior trends. For instance, identifying that workplace stress is heightened during specific time periods can lead to context-specific interventions, such as break reminders or task re-prioritization. This approach aligns with the principles of *context-aware computing*, emphasizing responsiveness to real-time factors to maximize the effectiveness of intervention [4].

While it is possible to conduct causal analysis generalized across all users, existing research demonstrates that causal relationships often exhibit significant variability between individuals due to differences in behavioral, psychological, and contextual factors [6]. This variability is consistent with our findings in Figure 5, which highlight the limitations of one-size-fits-all approaches. Instead, personalized interventions informed by individual-level causal analysis have shown greater potential to improve health outcomes [15]. By tailoring strategies to address specific causal pathways influencing behavior, personalized interventions can adapt dynamically to the unique needs and contexts of individual users, offering a more effective alternative to generalized approaches.

Building on the capabilities of **CausalCFF**, designing effective interventions requires the following considerations. First, interventions should be closely aligned with the specific causal relationships identified, ensuring relevance to the user's immediate context. Second, they need to adapt dynamically as user behaviors and stressors evolve. Third, minimizing user burden is crucial; interventions should provide concise, actionable recommendations rather than overwhelming users with excessive options. Finally, timing plays a key role; interventions delivered proactively (e.g., before entering high-stress contexts) are more likely to succeed. By addressing these factors, **CausalCFF**-based interventions can achieve greater impact and usability in real-world applications.

4.3 Limitations & Future Work

The suggested pipeline in this work has room for improvement. First, deterministic condition checking for CCM causal strength estimation [26] should be conducted, which was assumed to be true due to the large number of possible contextually filtered feature candidates. This is particularly challenging when dealing with a large number of pairs with potential causal relationships, which can also happen when balancing confounders in matching-based approaches. Previously, hypotheses on a causal scenario were made first, requiring researchers to conduct condition checking for a limited number of cases. The contextually filtered features in Figure 4 are not quite interpretable for users. Further research is needed on how to deliver these findings in a way that allows users to understand and utilize them in health management such as in Personal Informatics systems. Moreover, as mentioned in the previous section, we may extend the range of contextually filtered features by utilizing sequential rule mining when extracting features [14] to take the temporal order of events into account.

5 Conclusion

This study presents a novel methodology for establishing causal relationships between contextually filtered features extracted from mobile sensor data and human well-being. The approach advances the detection of high-level behavioral biomarkers associated with well-being, offering researchers a framework to design context-aware interventions that dynamically adapt to users' shifting contexts and behavioral patterns. Applied specifically to stress analysis, the method identifies contextual factors that significantly influence user stress levels. A key finding reveals that increased frequency of workplace visits—in contexts characterized by limited time spent at home—emerges as the most salient cause for heightened stress. This insight underscores the method's capacity to uncover granular, context-dependent drivers of well-being, thereby informing personalized intervention strategies.

Acknowledgments

This research was supported by the National Research Foundation of Korea NRF funded by Ministry of Science and ICT $NRF - 2022M3J6A1063021, 2022R1A2C2011536$

References

[1] Jumabek Alikhanov, Panyu Zhang, Youngtae Noh, and Hakil Kim. 2024. Design of Contextual Filtered Features for Better Smartphone-User Receptivity Prediction. *IEEE Internet of Things Journal* 11, 7 (2024), 11707–11722. https://doi.org/10.1109/JIOT.2023.3331715

[2] A. Bandura. 1986. Social Foundations of Thought and Action. https://api.semanticscholar.org/CorpusID:142519016

[3] B. E. L. Borja, S. C. Ibañez, and M. E. M. Ventura. 2020. *Time Series Analysis Handbook*. Asian Institute of Management. https://phdinds-aim.github.io/time_series_handbook/Preface/Preface.html

[4] Anind K Dey. 2001. Understanding and using context. *Personal and Ubiquitous Computing* 5, 1 (2001), 4–7. https://doi.org/10.1007/s007790170019

[5] Yunjo Han, Hyemin Lee, Kobiljon E. Toshnazarov, Youngtae Noh, and Uichin Lee. 2023. StressBal: Personalized Just-in-time Stress Intervention with Wearable and Phone Sensing. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers* (Cambridge, United Kingdom) *(UbiComp/ISWC '22 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 41–43. https://doi.org/10.1145/3544793.3560324

[6] Miguel A Hernán and James M Robins. 2020. *Causal Inference: What If.* Chapman & Hall/CRC. https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

[7] Duligur Ibeling and Thomas Icard. 2023. Comparing Causal Frameworks: Potential Outcomes, Structural Models, Graphs, and Abstractions. In *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Eds.), Vol. 36. Curran Associates, Inc., 80130–80141. https://proceedings.neurips.cc/paper_files/paper/2023/file/fd83f4e0dcaf1c64ea15bbb1695bb40f-Paper-Conference.pdf

[8] Gyuwon Jung, Sangjun Park, and Uichin Lee. 2024. DeepStress: Supporting Stressful Context Sensemaking in Personal Informatics Systems Using a Quasi-experimental Approach. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1000, 18 pages. https://doi.org/10.1145/3613904.3642766

[9] Gyuwon Jung, Sangjun Park, Eun-Yeol Ma, Heeyoung Kim, and Uichin Lee. 2024. Tutorial on Matching-based Causal Analysis of Human Behaviors Using Smartphone Sensor Data. *ACM Comput. Surv.* 56, 9, Article 236 (apr 2024), 33 pages. https://doi.org/10.1145/3648356

[10] Jaejeung Kim, Joonyoung Park, Hyunsoo Lee, Minsam Ko, and Uichin Lee. 2019. LocknType: Lockout Task Intervention for Discouraging Smartphone App Use. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300927

[11] Taewan Kim, Haesoo Kim, Ha Yeon Lee, Hwarang Goh, Shakhboz Abdigapporov, Mingon Jeong, Hyunsung Cho, Kyungsik Han, Youngtae Noh, Sung-Ju Lee, and Hwajung Hong. 2022. Prediction for Retrospection: Integrating Algorithmic Stress Prediction into Personal Informatics Systems for College Students' Mental Health. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 279, 20 pages. https://doi.org/10.1145/3491102.3517701

[12] Predrag Klasnja, Stephanie Smith, Nicholas J Seewald, Amy Lee, Kimberly Hall, Benjamin Luers, Eric B Hekler, and Susan A Murphy. 2019. Efficacy of contextually tailored suggestions for physical activity: A micro-randomized optimization trial of HeartSteps. *Annals of Behavioral Medicine* 53, 6 (2019), 573–582. https://doi.org/10.1093/abm/kay067

[13] Minsam Ko, Subin Yang, Joonwon Lee, Christian Heizmann, Jinyoung Jeong, Uichin Lee, Daehee Shin, Koji Yatani, Junehwa Song, and Kyong-Mee Chung. 2015. NUGU: A Group-based Intervention App for Improving Self-Regulation of Limiting Smartphone Use. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1235–1245. https://doi.org/10.1145/2675133.2675244

[14] Hansoo Lee, Auk Kim, SangWon Bae, and Uichin Lee. 2024. S-ADL: Exploring Smartphone-based Activities of Daily Living to Detect Blood Alcohol Concentration in a Controlled Environment. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1005, 25 pages. https://doi.org/10.1145/3613904.3642832

[15] Uichin Lee, Gyuwon Jung, Eun-Yeol Ma, Jin San Kim, Heepyung Kim, Jumabek Alikhanov, Youngtae Noh, and Heeyoung Kim. 2023. Toward Data-Driven Digital Therapeutics Analytics: Literature Review and Research Directions. *IEEE/CAA Journal of Automatica Sinica* 10, 1 (2023), 42–66. https://doi.org/10.1109/JAS.2023.123015

[16] Abhinav Mehrotra, Fani Tsapeli, Robert Hendley, and Mirco Musolesi. 2017. MyTraces: Investigating Correlation and Causation between Users' Emotional States and Mobile Phone Interaction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 83 (sep 2017), 21 pages. https://doi.org/10.1145/3130948

[17] Raul Montoliu, Jan Blom, and Daniel Gatica-Perez. 2013. Discovering Places of Interest in Everyday Life from Smartphone Data. *Multimedia Tools and Applications* 62, 1 (2013), 179–207. https://doi.org/10.1007/s11042-011-0982-z

[18] Adiba Orzikulova, Hyunsung Cho, Hye-Young Chung, Hwajung Hong, Uichin Lee, and Sung-Ju Lee. 2023. FinerMe: Examining App-level and Feature-level Interventions to Regulate Mobile Social Media Use. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 274 (oct 2023), 30 pages. https://doi.org/10.1145/3610065

[19] James O Prochaska and Carlo C DiClemente. 1983. Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology* 51, 3 (1983), 390–395. https://doi.org/10.1037/0022-006X.51.3.390

[20] Zhanna Sarsenbayeva, Gabriele Marini, Niels van Berkel, Chu Luo, Weiwei Jiang, Kangning Yang, Greg Wadley, Tilman Dingler, Vassilis Kostakos, and Jorge Goncalves. 2020. Does Smartphone Use Drive our Emotions or vice versa? A Causal Analysis. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376163

[21] Thomas Stütz, Thomas Kowar, Michael Kager, Martin Tiefengrabner, Markus Stuppner, Jens Blechert, Frank H. Wilhelm, and Simon Ginzinger. 2015. Smartphone Based Stress Prediction. In *User Modeling, Adaptation and Personalization*,

Francesco Ricci, Kalina Bontcheva, Owen Conlan, and Séamus Lawless (Eds.). Springer International Publishing, Cham, 240–251. https://doi.org/10.1007/978-3-319-20267-9_20

[22] George Sugihara, Robert May, Hao Ye, Chih hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. 2012. Detecting Causality in Complex Ecosystems. *Science* 338, 6106 (2012), 496–500. https://doi.org/10.1126/science.1227079

[23] Juho Sun, Sangkeun Park, Gyuwon Jung, Yong Jeong, Uichin Lee, Kyong-Mee Chung, Changseok Lee, Heewon Kim, Suhyon Ahn, Ahsan Khandoker, and Leontios Hadjileontiadis. 2020. BeActive: Encouraging Physical Activities with Just-in-time Health Intervention and Micro Financial Incentives. In *Proceedings of the 2020 Symposium on Emerging Research from Asia and on Asian Contexts and Cultures* (Honolulu, HI, USA) *(AsianCHI '20)*. Association for Computing Machinery, New York, NY, USA, 17–20. https://doi.org/10.1145/3391203.3391206

[24] Floris Takens. 1981. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980*, David Rand and Lai-Sang Young (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 366–381. https://doi.org/10.1007/BFb0091924

[25] Fani Tsapeli and Mirco Musolesi. 2015. Investigating Causality in Human Behavior from Smartphone Sensor Data: A Quasi-Experimental Approach. *EPJ Data Science* 4 (05 2015). https://doi.org/10.1140/epjds/s13688-015-0061-1

[26] Niels van Berkel, Simon Dennis, Michael Zyphur, Jinjing Li, Andrew Heathcote, and Vassilis Kostakos. 2021. Modeling interaction as a complex system. *Human–Computer Interaction* 36, 4 (2021), 279–305. https://doi.org/10.1080/07370024.2020.1715221

[27] Xuhai Xu, Prerna Chikersal, Afsaneh Doryab, Daniella K. Villalba, Janine M. Dutcher, Michael J. Tumminia, Tim Althoff, Sheldon Cohen, Kasey G. Creswell, J. David Creswell, Jennifer Mankoff, and Anind K. Dey. 2019. Leveraging Routine Behavior and Contextually-Filtered Features for Depression Detection among College Students. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 116 (sep 2019), 33 pages. https://doi.org/10.1145/3351274

[28] Panyu Zhang, Gyuwon Jung, Jumabek Alikhanov, Uzair Ahmed, and Uichin Lee. 2024. A Reproducible Stress Prediction Pipeline with Mobile Sensor Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 3, Article 143 (Sept. 2024), 35 pages. https://doi.org/10.1145/3678578