Utilizing Multi-Modal Personal Health Tracking and Health Affordances of the Built Environment

Clayton Feustel
Georgia Institute of Technology
feustel@gatech.edu

Lauren Wilcox
Georgia Institute of Technology
wilcox@gatech.edu

Submitted to the ACM DIS 2017 Workshop on People, Personal Data and the Built Environment

Abstract
In this workshop paper we describe our efforts toward the use of large-scale personal health tracking in developing a better understanding of the built environment and its impact on the health behavior of those dwelling in the space. As part of a larger mobile sensing study, we collected a wide array of health-related data from university students using phone sensors and Ecological Momentary Assessment. Utilizing this dataset, we are developing a process in which we map elements in the built environment to corresponding health behaviors that they afford. Ultimately, this mapping can lead to the creation of a generalized health score that can be applied at a building or site level. We are currently investigating how to structure and represent the data and models used to calculate this score—to enable personal and site-based reflection on these data.

Author Keywords
Built environment; personal health tracking

ACM Classification Keywords
H.5.m. Information interfaces and presentation: Miscellaneous;

Introduction

Thanks to the rapid adoption of smartphones and wearable devices, researchers interested in tracking an individual's health behavior are able to quickly, and fairly painlessly, get access to a wealth of high-quality sensor data. Many projects in the ubiquitous computing and HCI domains have used these sensor streams to inform tools that address peoples' health-related needs. Some have sought to improve health outcomes through reflecting personal health data to the user in consumer-friendly interfaces [1, 2]. While others, like the StudentLife study at Dartmouth College, attempt to utilize a mass collection of individual sensor data to understand health patterns at a community level [3]. Few systems that leverage these personal sensing modalities, however, have stepped outside the level of the individual when evaluating health, to consider the environmental context that affects behavior. Separating the concept of health from the individual and, instead, focusing on understanding it at an environmental level, provides invaluable information regarding the health behaviors encouraged or hindered by the specific environments that individuals occupy.

The benefit to mapping the attributes of the environment to health affordances has not escaped the notice of urban design and public health institutional bodies. In 2002, the Centers for Disease Control and Prevention held a workshop that pulled together an interdisciplinary team of experts, ranging from public policy to social marketing, to discuss the built environment's impact on public health. This workshop resulted in 37 questions that are intended to direct future research in the field [4]. From this list of research questions, Dannenberg et al. highlight selected topics which they propose as structuring elements for future research as well as hurdles that the field must overcome including *identifying exposure elements* and *defining guidelines.*

Using these themes as a framework, we explore how large-scale health tracking can be used to not only uncover relationships between the built environment and personal health but also understand how this environmental information can be utilized by systems to better inform the user of their health behavior.

Background

In 2016, we launched a research initiative at Georgia Tech, Campus Life, creating a database of continuous behavioral health and mental wellness data gathered from undergraduate and graduate students throughout their daily lives [5]. The Campus Life project collects two categories of data, passive sensor data and active self-reported data. The passive data streams are automatically collected at predetermined polling rates from the phone throughout the duration of the study and include accelerometry, GPS, ambient noise levels, application history, message metadata, call metadata, screen state, and fused activity data. Active self-report data is provided through Ecological Momentary Assessments (EMA) at semi-random intervals throughout the day and address mood, stress, and activity.

An initial deployment of the study was conducted in April of last year, with a larger deployment planned for the fall of 2017. This deployment resulted in the collection of 358 million health-related data points from 50 Georgia Tech undergraduate and graduate students. We plan on using this dataset to explore methods of describing the public health impacts of the built environment.

Goals of our Current Work

Identifying Exposure Elements. As described by Dannenberg et al., identifying exposure elements entails discovering and understanding the "best measures of the physical environment" that impact health. By utilizing daily tracked health markers of individuals in the Campus Life study, we aim to build a model that is able to map a relationship between environmental elements and the health behavior afforded by them.

Defining Guidelines. Guidelines can scaffold our understanding of the factors that play a role in determining the generalized health impact of an environment. Similar to urban design researchers' goal to create *walkability scores* by looking at factors such as residential density and retail to residential ratio [6], our objective is to discover the primary environmental factors that have the strongest impact on user's health behavior. Our analysis of the Campus Life data and the construction of the model to link exposure elements to behavioral health implications will be integral in outlining some of the factors most relevant to consider in examining the built environment's impact on an individual's behavioral health and mental wellness. Using the model, we will be able to identify the environmental artifacts that appear to have the largest impact on user's health behavior. We will use these artifacts as suggestions for potential guidelines in a general behavioral health score.

Behavioral Health Score. The creation of a behavioral health score will provide invaluable information to designers and users. Equally as important as creating the scoring process, is developing an understanding of how we frame this information to best support users in achieving their health goals.

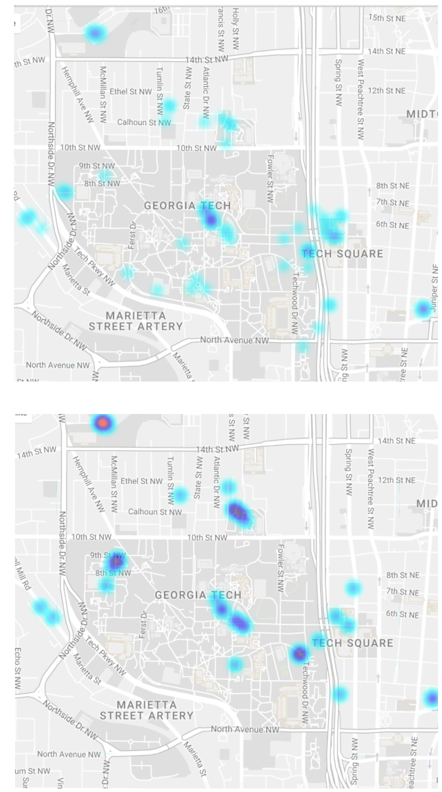

Figure 1. Two visualizations of active sensor data from CampusLife research participants, indicating where participants ate during the study period (top) and where they reported sleeping (bottom).

Proposed Approach

Identifying Exposure Elements

We will first identify all the locations on campus that enough participants have visited to provide a sufficient volume of data collected at each coordinate. We will then conduct an audit of the locations with both urban design and public health experts, taking them through an archetypal path traveled by a participant and labeling features that these experts note as artifacts that would impact health behavior. Expert observations will be marked at their geographical location and tagged according to a description of the artifact. A subset of duplicate artifacts will be removed from the dataset and reserved for testing. Tags will then be used with geographically situated participant behavioral data and supervised machine learning will be applied to construct a mapping between fluctuations in health behavior, as observed in the user data, and environmental artifacts. To evaluate our success in identifying health effects of exposure elements, we will test fit with the reserved environmental artifacts.

Defining Guidelines

With the model created from identifying exposure elements, we will select three artifacts that showed the strongest predictive ability. Using these three artifacts, we will create a scoring algorithm that produces a single value indicating the environment's health. The algorithm will take as input a set of artifacts and output a score. From the tagged dataset we will locate five Georgia Tech buildings that contain at least one artifact within a radius of the building's GPS coordinates. Feeding these buildings through the health score algorithm will produce a set of five scores. Finally, we will rank these buildings according to the returned health score.

To get a sense of how well the score captured the health affordances, we will compare the computational rankings to the rankings completed by the health experts. A panel of health experts (e.g. Georgia Tech clinicians) will review a visualization displaying the aggregate health data that was collected at each building. We will then conduct a think-aloud and have the clinicians rank the visualizations based on their perceptions of which datasets are healthiest. The observations from the think-aloud will be used to help explain errors in the algorithm and to evaluate how an automated approach may differ in its descriptive interpretation of the data.

Behavioral Health Score

A behavioral health score for buildings is of little use unless it can be situated in larger health systems to inform more intelligent applications and inform human decision-making and behavior. While we believe there are many use cases for a behavioral health score, including systems for administrators and architects, we want to focus on how environmental health information can be reflected back to users. As data about the environment is rarely presented in health tracking applications, there are many things that we must consider as designers to ensure that the information is adequately displayed.

The first question that must be addressed is how we allow users to interact with the environment data. One possibility is to simply present the score of each building to the users in an information card. One attractive aspect of such a design is that it provides users with the most explicit display of how the environment plays a role in their health. A potential downfall of this design, however, is the possibility that users may misinterpret the intention of the health score. For example, a user who sees that a specific building strongly discourages physical activity, may try to avoid the building. A building that discourages physical activity is not inherently bad; the building simply is not an optimal location for participating in activities in pursuit of reaching an individual's physical activity goal.

Another option is to allow the user to explore the environment data through controlled queries. For example, if a user is looking for a location to calm down, she may query the application for buildings that strongly afford calm behavior. A querying system allows users to easily understand the relevant environmental context for a specific health behavior, but may hinder the ability to see broader relationships between personal health and the environment.

How we display the environmental health data is another aspect that we must consider when designing a user-centered system. There is an inherently geographical basis to visualizing the built environment's health affordances. Health data is often shown through a temporal frame, however. Finding a way to combine these disparate data streams will be a difficult but necessary hurdle for incorporating the built environment into user health applications. With these design considerations in mind, we will conduct a participatory design study with Campus Life participants to create an effective interface that integrates traditional health data with environmental health behavior data.

Impact and Future Work

Ultimately, we aim to forge advancements in behavioral and environmental sensing, behavioral modeling, personal health informatics, and information abstraction, while improving our understanding of the relationship between built environments and human health. As such, we wish to connect with researchers in the design, human–computer interaction, public health and ubiquitous computing communities, who are exploring this emerging area of research. Collectively, our efforts can contribute powerful additions to traditional health interventions while advancing data visualization, abstraction, and reflection techniques. These techniques can evolve to incorporate representations of health affordances of built environments. For example, an ever-present concern for research when designing a visualization system is information overload [7]. Trimming and highlighting the data based on the health affordances of the environment—by showing only the health measures that are most afforded by the user's current location—is one way an application could reduce cognitive load and maximize information relevancy.

There are three primary contributions to the built environments and design communities: 1) a method of identifying artifacts in the environment that impact health through mobile data-collection methods, 2) suggested health artifacts that may be useful in defining a general understanding of an environment's health, and 3) an overview of potential design considerations for systems intending to incorporate environmental-based health data. At the workshop, we are hoping to get reactions from researchers about the use of multi-modal personal health tracking data as a means of determining health affordances. We are also asking for feedback on the potential benefits and pitfalls of designing applications that integrate environmental health affordance data with traditional personal health metrics.

References

1. Bentley, F., Tollmar, K., Stephenson, P., Levy, L., Jones, B., Robertson, S., … Wilson, J. (2013). Health Mashups. ACM Transactions on Computer-Human Interaction, 20(5), 1–27. https://doi.org/10.1145/2503823ACM. How to Classify Works Using ACM's Computing Classification System. 2014. Retrieved August 22, 2014 from http://www.acm.org/class/how_to_use.html.

2. Brownson, Ross C., Hoehner, C., Day, K., Forsyth, A., S. J. (2010). Measuring the Built Environment for Physical Activity: State of the Science. *American Journal of Preventive Medicine*, *36*, 53. https://doi.org/10.1016/j.amepre.2009.01.005.Measuring

3. StudentLife Study. (n.d.). Retrieved April 26, 2017, from http://studentlife.cs.dartmouth.edu/

4. Dannenberg, A. L., Jackson, R. J., Frumkin, H., Schieber, R. A., Pratt, M., Kochtitzky, C., & Tilson, H. H. (2003). The Impact of Community Design and Land-Use Choices on Public Health: A Scientific Research Agenda. *American Journal of Public Health*, *93*(9), 1500–1508. https://doi.org/10.2105/AJPH.93.9.1500

5. Georgia Tech Campus Life. (n.d.). Retrieved April 26, 2017, from http://www.quantifiedcampus.gatech.edu/

6. Frank, L. D., Sallis, J. F., Saelens, B. E., Leary, L., Cain, K., Conway, T. L., & Hess, P. M. (2010). The development of a walkability index: application to the Neighborhood Quality of Life Study. *British Journal of Sports Medicine*, *44*(13), 924–933. Retrieved from http://bjsm.bmj.com.ezproxy.library.ubc.ca/content/44/13/924

7. Brownson, Ross C., Hoehner, C., Day, K., Forsyth, A., S. J. (2010). Measuring the Built Environment for Physical Activity: State of the Science. *American Journal of Preventive Medicine*, *36*, 53. https://doi.org/10.1016/j.amepre.2009.01.005.Measuring