Descriptive Analysis of Physical Activity Conversations on Twitter

Abstract
This paper explores how people are using Twitter.com to manage and share information about health-promoting physical activity. We analyzed archived posts, called “tweets”, from Twitter.com to learn about the range, patterns, and captured metadata associated with muscle-strengthening, aerobic, and flexibility-enhancing physical activities. The content analysis describes how people are using Twitter to post about their health-related fitness activities. These findings can support the design of supportive tools and applications connected with the social media platform.

Keywords
Twitter, social networks, Technology-Mediated Social Participation, micro blog, physical activity

ACM Classification Keywords
J.3 Life and Medical Sciences (Health), J.4 Social and Behavioral Sciences (Sociology)

General Terms
Human Factors

Introduction
Chronic non-communicable diseases (NCDs), such as diabetes and cancer, account for almost 60% of deaths worldwide [24]. Unhealthy lifestyle behaviors contribute significantly to the high prevalence of NCDs [23].
Governments are trying to counter this trend by promoting healthy behaviors, such as physical activity, that are known to reduce the risk of a range of health conditions, including heart disease, stroke, diabetes, hypertension, cancer, and depression [23].

With levels of physical inactivity rising globally, the challenge is to elicit participation in a largely sedentary population [23]. The health benefits of physical activity are only realized with regular, moderate-to-vigorous intensity workouts [18]. Leveraging social support can serve as a mechanism for maintaining this level of commitment and motivation to healthy behavior [15]. People frequently turn to their informal peer networks to ask for health advice and share their experiences [3]. The emergence of web-based social networking tools (e.g. Facebook.com and Twitter.com), allow people to make their activities visible and thus available to one another to seek expertise and social support. This paper explores the potential of a system like Twitter to aid people in drawing upon one another for health-related fitness.

**Background**

Over the past few years, the use of internet-based social software (e.g. online communities, blogs, wikis, and social networking tools) has grown substantially. Use of social networking sites nearly doubled from 35% of Internet users in 2008 to 61% in 2010, with the fastest rate of growth in adults 74 and older [25]. A growing trend in this space is the explosion in the use of Twitter. Eight percent of American adults use Twitter to share brief status updates (‘tweets’) ranging from observations about their personal and professional lives to their location [20].

At the forefront of this growth, people have begun to use social software to exchange health information with peers [6,19]. For example, the social networking tool PatientsLikeMe.com encourages users to share personal experiences about specific conditions, such as Multiple sclerosis [9]. A growing number of studies have investigated the use of social software for accessing health information ranging from topics such as ‘Cancer 2.0’ [7] to weight-loss in online communities [12].

Many people now use social software more often to obtain advice from other patients than to obtain emotional support [19]. Yet, although one-third of adults accessed social media related to health in 2009, few posted about their own health [6]. Could the recent growth in Twitter use for posting personal updates be changing this landscape? Although Twitter shows a fast rate of growth in use, [20] few studies have reported on how this platform influences social interaction about health. Thus, Twitter presents an interesting, yet underexplored tool for health-promoting activities.

In this work, we investigate how people are using Twitter for health-related interaction. Because use of the Internet for exercise and fitness information is one of the most prevalent and fastest areas of growth [6] we describe how people tweet about physical activity. This description will enhance our understanding of individuals’ online interactions about health and inform the design of tools that leverage Twitter to support people’s efforts to adopt healthy lifestyles.

**Methods**

We analyzed archived tweets from Twitter.com using a public API to query the stream of status messages. Specifically, we used the "Spritzer" stream that
provides a sample of 1% of the approximately 95 million status messages that are made every day [21]. In total, we sampled 15,000 messages, collected every 4 hours, from January 1 – January 10, 2011.

We designed our query to filter based on keywords that reflect modes of exercise as defined within the recommended guidelines for the US population: aerobic, muscle-strengthening, and flexibility-enhancing physical activities [10,18]. Respectively, we used the following terms for each of the three categories of physical activity: “elliptical”, “weight lifting”, and “Pilates”. The keywords reflect concepts with perceived low level of ambiguity, while enabling us to analyze messages related to all three types of physical activity recommended by the guidelines. To account for misspellings and alternative phrasing of the concepts, we also included “eliptical”, “weightlifting”, and “lifting weights” as part of our query.

**Automated content analysis.** We focused our analysis to identify tweets that incorporated URLs, hashtags, retweets, @replies, and @mentions. In addition, we extracted the source of each tweet to analyze how it was created (e.g., using the Twitter website, a mobile application, etc.).

**Manual analysis.** To validate the relevance of our selected keywords and to gain deeper understanding of the content of fitness-related tweets, we sampled 100 tweets from each of the three designated categories—elliptical, weight lifting, Pilates—and analyzed their content for high-level themes.

Our data is based on publicly available tweets from non-protected Twitter accounts. Tweets created by protected accounts and all personal direct messages are not available via the Streaming API.

**Results**

*Automated content analysis.* Of the 15,000 tweets sampled, 74% made reference to outside content or other users. More than 10% of the total posts included a question of some kind (Table 1).

<table>
<thead>
<tr>
<th>Posts with...</th>
<th>Tweet Count</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any linked reference</td>
<td>11,139</td>
<td>74%</td>
</tr>
<tr>
<td>URL</td>
<td>6,732</td>
<td>45%</td>
</tr>
<tr>
<td>@mention</td>
<td>4,027</td>
<td>27%</td>
</tr>
<tr>
<td>Hashtags</td>
<td>2,137</td>
<td>14%</td>
</tr>
<tr>
<td>@replies</td>
<td>1,989</td>
<td>13%</td>
</tr>
<tr>
<td>Question</td>
<td>1,439</td>
<td>10%</td>
</tr>
<tr>
<td>Re-tweet</td>
<td>998</td>
<td>7%</td>
</tr>
</tbody>
</table>

**Table 1:** Prevalence of tweets referencing outside content/users

Users posted messages to Twitter from a variety of sources. Some users posted directly on the Twitter website, while others used a mobile phone or computer-based application. A summary of how users submitted fitness-related tweets is outlined in Table 2.

<table>
<thead>
<tr>
<th>Source</th>
<th>Tweet Count</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web</td>
<td>7,152</td>
<td>48%</td>
</tr>
<tr>
<td>Mobile</td>
<td>4,483</td>
<td>30%</td>
</tr>
<tr>
<td>Multiple</td>
<td>2,532</td>
<td>17%</td>
</tr>
<tr>
<td>Computer-based app</td>
<td>58</td>
<td>&lt;1%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>14,225</strong></td>
<td><strong>95%</strong></td>
</tr>
</tbody>
</table>

**Table 2:** Distribution of tweets by type of source
Evidence of Exercise
"I just dominated the elliptical. This weight is history. Who’s ready to be foxy by summer? This gal...!"

Plans for Exercise
"Weight Lifting tomorrow after school. Pre-season officially begins. #WhatADread"

General Attitude
"I love elliptical machines!" "Pilates!"

Request
"Where can I go to rent pro-fitness gear? Don’t wanna buy but I need an elliptical for 4-6 months for an exec."

Advertisement
"Order your 2011 model Yowza elliptical machine today and Receive free shipping!"

Off-Topic
"surprised myself by learning how SVG elliptical arcs work"

Figure 1. Sample tweets for themes from manual review

Manual Review. Validation of a selected subset of the sample is shown in Table 3. Relevant tweets were defined as content directly about the targeted physical activities. Tweets were not relevant if the content was off-topic regarding the selected category. Figure 1 shows sample quotes for themes from manual review. Of the relevant tweets, at least one-third of elliptical and Pilates and more than half of the weight lifting tweets had content referring to an advertisement for a product or service (Table 3).

<table>
<thead>
<tr>
<th>Category</th>
<th>Relevant</th>
<th>Advertise</th>
<th>Non-advertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elliptical</td>
<td>97%</td>
<td>35%</td>
<td>62%</td>
</tr>
<tr>
<td>Weight Lifting</td>
<td>89%</td>
<td>53%</td>
<td>36%</td>
</tr>
<tr>
<td>Pilates</td>
<td>90%</td>
<td>30%</td>
<td>60%</td>
</tr>
</tbody>
</table>

Table 3: Percentage of relevant tweets by physical activity

Of the relevant, non-advertising tweets, the content primarily reflected evidence or reports of physical activity, announcements for planned physical activity, or general attitudes towards the fitness categories. Additionally, people used Twitter to make requests, provide feedback to others, and make recommendations.

Discussion
Our evaluation of the content of tweets showed a high prevalence of references to outside content and users. The use of @user through mentions and replies is comparable with other studies on Twitter in general, but the percentage of URLs in our sample is substantially higher [2,13]. This reflects in part the limited amount of content that can be expressed in 140 characters, but also suggests that Twitter is frequently used to connect users to other content on the web.

With more than 30% of tweets originating from mobile devices, Twitter users show a particular affinity for using mobile devices to interact about fitness. The number of mobile-based messages is likely higher because 17% of the sample were sent via applications (e.g., tweetDeck) that operate over multiple platforms (e.g., web, computers, smartphones, email). The prevalence of mobile-based tweets reflects a broad trend that wireless access is associated with deep engagement and participation in online interaction [6].

In 2010, Twitter enhanced its API and website so posts can include location information as specific as exact GPS coordinates or a general area like a neighborhood or town [22]. With the rapid growth of mobile devices equipped with GPS, the high prevalence of mobile posts means that fitness-related tweets will increasingly contain location information. This information opens up possibilities for the development of tools that leverage Twitter to provide users with location-based tips, recommendations, and other information that can support their efforts to adopt healthy lifestyles.

Among manually analyzed tweets, the two most frequent categories were posts in which users reported (1) evidence of or (2) plans for exercising. Although these posts are probably intended to provide their social network with an update on what they have been doing, the prevalence of these types of posts suggest that Twitter might also be a promising platform for leveraging social support to motivate health behavior change. For example, location information could be used to create ad hoc social groups of individuals within
a certain area who are pursuing similar health goals. Prior work on diabetes [8] and physical activity [4] has shown that sharing information about health-related activities provides powerful support for healthy habits.

Our analysis also revealed that Twitter is used to ask fitness-related questions and provide recommendations and resources. For example, users sometimes shared links to the workouts they created (e.g., "first elliptical workout ever http://bit.ly/..."). Twitter’s retweeting and @mention facilities, which enable wide broadcasts of such advice, make Twitter well-suited for easy exchange of expertise about fitness and health.

There are several limitations with the data gathered for this study. The sample was taken over 10 days during January 2011 and might not be representative due to day of week, time of day, or time of year. Moreover, the sample might reflect the fact that January is typically a time when people discuss personal resolutions relating to fitness and weight loss. Finally, selection bias in the specific terms used to identify tweets might have excluded relevant content.

**Conclusion**

Even with a limited set of keywords, it is clear that Twitter is an active forum for sharing personal health and exercise information. Use of social media for health-related fitness is likely to rise as usage patterns increase across age groups in the future. Recognizing this, governmental strategies to promote physical activity include specific aims to expand their use of media to social networking tools. Our findings also suggest a strong mobile presence among Twitter users. With mobile broadband subscriptions expected to hit one billion mark in 2011 [5], Twitter has the potential to be a powerful platform for innovative, mobile-based applications that promote health activities.

**Acknowledgements**

This work was supported in part by the NIH National Library of Medicine Biomedical and Health Informatics Training Program, Grant Number 5 T15 LM007442.

**References**


[8] Franklin, V.L., Waller, A., Pagliari, C., Greene, S.A. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with...

[9] Frost, J.H., Massagli, M.P. Social uses of personal health information within PatientsLikeMe, an online patient community: What can happen when patients have access to one another’s data. *J Med Internet Res* 10, 3 (2008), e15, PMID:18504244.


