Understanding and Utilizing the Influence of Social Networks on Health Care in YesiWell

Dejing Dou, Daniel Lowd, Jessica Greene, Brigitte Piniewski, Ruoming Jin
Xintao Wu, David Kil, and Frances Shin

December 15, 2010

Motivation: Two thirds of the US population are now overweight or obese. This incurs significant health risks and financial costs to society. Obesity and overweight, although multifactorial, can largely be explained by the social and cultural spread of poor health habits such as lack of exercise, smoking, fast food consumption, and alcoholism spread through communities, introducing and reinforcing harmful behaviors [1]. By the time behaviors develop into diseases and individuals seek medical help, it is often too late to reverse the chronic poor health outcomes. Traditionally, support groups and other social reinforcement approaches have been popular and effective in dealing with unhealthy behaviors including overweight. Of the factors associated with sustained weight loss one of the most important is continued intervention with frequent social contacts.

Recent advances in mobile technology and online social networks provide new opportunities to support healthy behaviors through lifestyle monitoring and online communities. The YesiWell project[1] uses both approaches together to help people maintain active lifestyles and lose weight. Though still in the pilot stages, YesiWell already has a wealth of data on the health statistics, activity levels, and social network activity of a large group of users. YesiWell is expected to have up to 2 million users in next couple of years. The YesiWell data specifically include three dimensions: i) physical activities such as the aerobic steps, regular steps, speed, and duration. Each user carries a sensor called *hPod* which can do physical activity tracking as well as aggregation of weight and blood pressure data. ii) biometrics and biomarkers such as serum creatinine, total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL) etc. iii) social activities tracked through user-generated social networks, and provides social support, discussions, and competitions.

Background: To exploit the full potential of wellness systems and social networks such as Yesi-Well, it will be necessary to develop new methods in data mining, social network analysis, knowledge representation, and privacy that can handle the complexities of data that is high-dimensional, large scale, temporal, social, and very sensitive. The authors of this white paper propose a novel framework to address this need. The research team include the experts in data mining, social networks, health policy, ontology, privacy, and health informatics. The first two authors are potential participants of the workshop if the paper is invited. Dejing Dou is an Associate Professor in Computer and Information Science (CIS) at the University of Oregon. Professor Dou is an expert in ontology-based data mining and data integration. He is the director of the Advanced Integration and Mining Laboratory. Professor Dou is specially interested in applying Semantic Web ontologies and data mining into biomedical

---

[1] http://www.yesiwell.com/

informatics and health informatics. He is currently the PI of a NIH funded four year R01 project titled with *Neural ElectroMagnetic Ontologies (NEMO): ERP Knowledge Representation and Integration*. Since the spring of 2010, Professor Dou has been collaborating with Dr. Brigitte Piniewski (PI of YesiWell), Prof. Jessica Greene (expert in health policy), Mr. David Kil (Chief Scientist of Yesi-Well) for data warehousing in YesiWell. Daniel Lowd is an Assistant Professor in the Department of Computer and Information Science at the University of Oregon. His research covers a range of topics in statistical machine learning, including statistical relational representations, unifying learning and inference, and adversarial machine learning applications (e.g., spam filtering). Prof. Lowd have been collaborating with Prof. Dou, Prof. Ruoming Jin (expert in social network mining), and Prof. Xintao Wu (expert in privacy preserving mining) for YesiWell data mining.

Vision: Our vision is to develop new methods in a variety of areas, building on the current state-of-the-art. We will utilize temporal Granger causality and dynamic Bayesian networks to study the influence of social networks on health care. We will design a formal semantic model with Semantic Web ontologies including common concepts and mappings for health behaviors, biomarker measures, and social activities, and a sophisticated way to reuse those ontologies to facilitate data mining and social network analysis. We will also develop novel differential privacy preserving techniques to provide strict privacy guarantees during the process of collecting, querying, analyzing, and mining such complex health data with social networks. Finally, we will evaluate our data mining algorithms both by the quantitative measures and the tailored communication from the social science point of view. Though we will focus on mining and modeling the real-world data from the YesiWell environment, the approaches, methodology, and results developed and discovered in the projects can be applied in any other emerging healthcare social networks. Given the widespread use of personnel health systems and the research on the influence of social networks, development of such data mining algorithms, social network design, ontologies, and privacy preserving methods is both timely and significant.

Internet-based health interventions hold particular promise for studying the impact of social networks on health care, especially for tailored communication. While Internet-based tailored health communications are widely used, the research in this area is still quite limited. Our research will help overcome these limitation by leveraging sophisticated data mining and ontology techniques. A further benefit will be the support for mining other cross-dimensional and complex network data. The prospect of data mining from different dimensional and large scale network data is exciting, though challenging for a variety of reasons. Many researchers would agree that multi-dimensional mining and social network analysis is one of the major aims for advancement of health informatics.

Also, understanding the influence of social networks on health activities could have clinical, as well as basic research, applications. For example, using the YesiWell data, will these new data mining tools help separate the clinically relevant co-occurrences that distinguish weight loss that is health promoting from weight loss that is malignantly disruptive and not health promoting as it results in a much higher risk of future weight gain? Clinically, it is apparent that many of our patients with the most aggressive and prolonged approach to dieting are often also the heaviest. Distinguishing between the co-occurrences that produce divergent health outcomes may be relatively challenging using institutional data. High yield (high predictive capacity) data from individuals and households (such as collected in the YesiWell system) may provide the powerful raw data to answer these questions.

References

[1] N. Christakis and J. Fowler. The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, 357(4), 2007.