Multidimensional Analysis of Heterogeneous Relationships in Internet of Things

THE UNIVERSITY of ADELAIDE

A dissertation submitted in fulfillment

of the requirements for the degree of

Doctor of Philosophy

in

Computer Science

Lina Yao

Supervisors: A/Prof. Michael Sheng

Prof. Anne H. H. Ngu

2014

ORIGINALITY STATEMENT

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the Universitys digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Lina Yao

January 10, 2014

To my parents,

my husband and my daughter

ACKNOWLEDGMENTS

I would like to thank all the people who have given me tremendous support and help to make this thesis happen.

First of all, I am deeply grateful to my supervisor, A/Prof. Michael Sheng, for his generous time and devotion on supervision and guidance during my PhD study. He instilled a thirst for excellence in me, taught me how to do high-quality research, and helped me think independently and creatively. In many ways, he has set up a great example for me: his passion on research and life, his patience in guiding students, and his diligence. I always feel lucky to have A/Prof. Michael Sheng as my supervisor, for all the earnest encouragements, and the strongest supports in every means.

I sincerely thank Prof. Anne H.H. Ngu and Assistant Professor Byron Gao, who have given me many valuable suggestions and insightful discussions in my research. I would like to thank Prof. David Suter who has been supporting me in many ways, from scholarship applications to fellowship application. I have also benefited a lot from his course on Computer Vision, which opens a fascinating research world for me.

I would like to thank my co-authors: Quan Z. Sheng, Anne H.H. Ngu, Byron Gao, Zakaria Marmarr, Xue Li, Talal Noor, Claudia Szabo, Aviv Segev, Rui Zeng for their productive and enjoyable collaborations.

I would like to also thank my friends for providing support and friendship whenever I needed. Especially thank my two dearest friends, Bo Zhao, for his unconditional support all the time, Lijuan Wang, who has kindly given me the precious help during my hardest time.

Finally and above all, I owe my deepest gratitude to my family. I want to thank my mother Junling Liu and my father Guangquan Yao for their endless and unreserved love, who have been encouraging and supporting me all the time. Their love has provided my inspiration and has been my driving force. Without their patience and love I would have not survived the difficult times. I want to thank my husband, Tao Wang, for his love, understanding, patience, and support at every moment. I also want to thank my new born daughter Alicia, who is a miracle in my life and gives me ongoing courage to go through the challenging journey. This thesis is dedicated to them.

ABSTRACT OF THE DISSERTATION

Multidimensional Analysis of Heterogeneous Relationships in Internet of Things

by

Lina Yao

Doctor of Philosophy in Computer Science

The University of Adelaide, 2014

The Internet of Things (IoT) describes the evolution from systems linking digital information to systems relating digital information to real world physical items. In IoT, real-world objects are being connected to the Internet and interacted with traditional web entities (images, videos, texts and people in social networks), which complicates the relationships amongst heterogeneous entities. While it is well understood that IoT offers numerous opportunities and benefits, it also presents significant technical challenges. Among them, effective and efficient management of swift growing ubiquitous things in the boosting IoT is a fundamental challenge facing society and research community today. The essential prerequisite is efficient acknowledging and analysis of *heterogeneous relationships* existing in Internet of Things, and with it, many research challenges can be pushed forward.

In this thesis, we have systematically explored the principles and developed a series of methodologies of mining and modeling *three* types of heterogeneous relationships existing in Internet of Things.

Implicit mutual correlations of things. We study the problem of discovering implicit correlations among ubiquitous things, which have unique and challenging characteristics (e.g., no uniform features, diverse, and dynamic). We explore explicit user

interactions with things and develop two graph-based models, namely T-DisCor and T-DisCor+, for discovering latent correlations among things over graphs induced from thing usage events via modeling the contextual information and entity relationships respectively. We demonstrate how the discovered correlations of things can contribute to solving a number of important applications on things management. In particular, we develop an effective and flexible feature-based method for annotating things based on the outcome of our proposed approach.

Relationship between things and its attributes. We investigate the problem of exploiting relationships between ubiquitous things and their different attributes, we propose multiple approaches:

- Thing's semantic labels. We propose a semi-supervised learning framework for multi-label classification, which fully considers objects mutual affinity, semantic label correlation, coupled by thing-label assignment relationships for effective semi-supervised classification.

- Thing's social tags. We explore social tags for web object classification by developing a tag-centric, unified and discriminative classification framework. We not only use social tags conveying partial and latent information about the web objects as a novel evidence to facilitate classifying objects on the web, but also exploit the relative information among tags.

- Thing's availability. We propose an algorithm to estimate the availability of things so that those with higher availability can be selected and delivered in applications such as things composition.

Interactive relationship between people and things. Things recommendation is a crucial step for promoting and taking full advantage of Internet of Things, where it benefits the individuals, businesses and society on a daily basis in terms of two main aspects. On the one hand, it can deliver relevant things (things have similar functionalities that users might need) to users based on users preference and interest. On the other hand, it can also serve to optimize the time and cost of using IoT in a particular situation. Physical things in reality have multiple unique attributes. We propose two novel approaches to explore things recommendation problem.

- We propose a probabilistic matrix factorization based joint model to address things recommendation problem in Internet of Things. We fuse information from users social networks and things correlation networks, by learning shared latent factors stemming from the probabilistic matrix factorization on three matrices, namely users relations, things correlations, and observable things usage interactions.

- We propose a novel approach that unifies collaborative filtering and content-based recommendations. In particular, our approach considers simultaneously both rating data (e.g., QoS) and semantic content data (e.g., functionalities) of services offered by things using a three-way aspect model. Unobservable user preferences are represented by introducing a set of latent variables, which is statistically estimated.

For each of model and approach, we have conducted extensive experiments to validate and evaluate it using either public data set or datasets generated from an environment we built.

The studies presented in this dissertation lead to a series of models and real-world applications, including things classification, things clustering and things recommendation. Some open research directions are also discussed in the dissertation.

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables