

UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL AND APPLIED SCIENCES

Electronics and Computer Science

Using Linked Data in Purposive Social Networks

by

Priyanka Singh

Thesis for the degree of Doctor of Philosophy

February 2015

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCES

Electronics and Computer Science

Doctor of Philosophy

USING LINKED DATA IN PURPOSEFUL SOCIAL NETWORKS

by Priyanka Singh

This work is all about ...

Contents

Declaration of Authorship	xiii
Acknowledgements	xv
Nomenclature	xvii
1 Introduction	1
1.1 Overview	1
1.2 Research Challenge	1
1.3 Research Contribution	1
1.4 Structure of Thesis	1
2 Background	3
2.1 Emergence of Social Web	5
2.1.1 Web 2.0 and Social Networks	5
2.1.2 Content-specific Social Networking Services	5
2.2 Collective Intelligence and Crowdsourcing	5
2.2.1 Crowdsourcing Services	5
2.2.1.1 Wiki	5
2.2.1.2 Forums and Messaging Boards	5
2.2.1.3 Q&A Services	5
2.2.2 Human Computation	5
2.2.3 User Recommendation System	5
2.3 Semantic Web and Linked Data	5
2.3.1 Social Semantic Technology	5
2.3.1.1 FOAF	5
2.3.1.2 SIOC	5
2.3.1.3 OPO	5
2.3.2 Linked Data Technology	5
2.3.2.1 Linking the Datasets	5
2.3.3 Open Linked Data and Graph	5
2.4 Semantic Search	5
2.4.1 Keyword Disambiguation	5
2.4.2 Concept Mapping	5
2.4.3 SPARQL Queries	5
3 Purposive Social Network	7
3.1 What is Purposive Social Network	9

3.1.1	Information Based Community	9
3.1.2	Interest Based Community	9
3.1.3	Expert Based Community	9
3.1.4	Location Based Community	9
3.2	Characteristics of Purposive Social Network	9
3.2.1	Community Size	9
3.2.2	Focused Interest	9
3.2.3	Direct Communication	9
3.2.4	Active Participation	9
3.2.5	Short Lifespan	9
3.2.6	Strong Incentive	9
3.3	Benefits of Purposive Social Network	9
3.3.1	Information Exchange and Self-interest	9
3.3.2	Symbiotic Relation and Social Exchange	9
3.3.3	Social Recognition and Personal Satisfaction	9
3.3.4	Recommendation System	9
3.3.5	Expert Finder	9
3.4	Crowdsourcing in Purposive Social Network	9
3.4.1	Recruiting and Retaining Users	9
3.4.2	Incentive Model	9
3.4.3	Quality Control	9
3.4.4	Search and Discovery of Quality Content	9
3.5	Use of Semantic Web and Linked Data in Purposive Social Network	9
3.5.1	Structured Data	9
3.5.2	Linking People to People and People to Data	9
3.5.3	Multidimensional Network and Graph	9
3.5.4	Integrated Knowledgebase	9
3.5.5	Smart Query and Search	9
3.5.6	Social Network Analysis	9
4	System Design and Methodology	11
4.1	Data Mining	12
4.1.1	StackOverflow Data Mining	12
4.1.2	Reddit Data Mining	12
4.2	Data Structuring	12
4.2.1	Data Cleaning	12
4.2.2	Converting into N-Triples	12
4.2.3	Linking the Data	12
4.3	Keyword Disambiguation	12
4.3.1	OpenCalais Service	12
4.3.2	DBpedia Spotlight Service	12
4.4	Concept Mapping	12
4.4.1	Keyword Matrix	12
4.4.2	Document Term Matrix	12
4.5	Search and Query	12
4.5.1	Database Design	12
4.5.2	Searching Answers of Unanswered Questions	12

4.5.3	Expert Finder	12
5	Purposive Social Network Analysis	13
5.1	Network Linkage and Social Ties	13
5.2	Role of Individual Actors	13
5.3	Incentive Design	13
5.4	Quality Control	13
5.5	Linked Data Graph	13
5.6	Discourse Analysis	13
6	Evaluation	15
6.1	Experiment Design	15
6.1.1	Calculating Sample Size	15
6.1.2	Selecting Questions	15
6.2	Data Collection	15
6.2.1	Questionnaire Design	15
6.2.2	Dataset Frequency Distribution	15
6.3	Results	15
6.3.1	Keywords T-Test	15
6.3.2	Q&A Correlation Test	15
6.3.3	Summary	15
7	Conclusions	17
7.1	Summary of Results	17
7.2	Limitation of System	17
7.3	Future Work	17
A	Stuff	19

List of Figures

List of Tables

Declaration of Authorship

I, **Priyanka Singh** , declare that the thesis entitled *Using Linked Data in Purposive Social Networks* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as: (?), (?) and (?)

Signed:.....

Date:.....

Acknowledgements

Thanks to no one.

Nomenclature

w The weight vector

Chapter 1

Introduction

1.1 Overview

1.2 Research Challenge

1.3 Research Contribution

1.4 Structure of Thesis

Chapter 2

Background

Hello

2.1 Emergence of Social Web

2.1.1 Web 2.0 and Social Networks

2.1.2 Content-specific Social Networking Services

2.2 Collective Intelligence and Crowdsourcing

2.2.1 Crowdsourcing Services

2.2.1.1 Wiki

2.2.1.2 Forums and Messaging Boards

2.2.1.3 Q&A Services

2.2.2 Human Computation

2.2.3 User Recommendation System

2.3 Semantic Web and Linked Data

2.3.1 Social Semantic Technology

2.3.1.1 FOAF

2.3.1.2 SIOC

2.3.1.3 OPO

2.3.2 Linked Data Technology

2.3.2.1 Linking the Datasets

2.3.3 Open Linked Data and Graph

2.4 Semantic Search

2.4.1 Keyword Disambiguation

2.4.2 Concept Mapping

2.4.3 SPARQL Queries

Chapter 3

Purposive Social Network

Hello

3.1 What is Purposive Social Network

3.1.1 Information Based Community

3.1.2 Interest Based Community

3.1.3 Expert Based Community

3.1.4 Location Based Community

3.2 Characteristics of Purposive Social Network

3.2.1 Community Size

3.2.2 Focused Interest

3.2.3 Direct Communication

3.2.4 Active Participation

3.2.5 Short Lifespan

3.2.6 Strong Incentive

3.3 Benefits of Purposive Social Network

3.3.1 Information Exchange and Self-interest

3.3.2 Symbiotic Relation and Social Exchange

3.3.3 Social Recognition and Personal Satisfaction

3.3.4 Recommendation System

3.3.5 Expert Finder

3.4 Crowdsourcing in Purposive Social Network

3.4.1 Recruiting and Retaining Users

3.4.2 Incentive Model

3.4.3 Quality Control

3.4.4 Search and Discovery of Quality Content

Chapter 4

System Design and Methodology

Hello

4.1 Data Mining

4.1.1 StackOverflow Data Mining

4.1.2 Reddit Data Mining

4.2 Data Structuring

4.2.1 Data Cleaning

4.2.2 Converting into N-Triples

4.2.3 Linking the Data

4.3 Keyword Disambiguation

4.3.1 OpenCalais Service

4.3.2 DBpedia Spotlight Service

4.4 Concept Mapping

4.4.1 Keyword Matrix

4.4.2 Document Term Matrix

4.5 Search and Query

4.5.1 Database Design

4.5.2 Searching Answers of Unanswered Questions

4.5.3 Expert Finder

Chapter 5

Purposive Social Network Analysis

Hello

5.1 Network Linkage and Social Ties

5.2 Role of Individual Actors

5.3 Incentive Design

5.4 Quality Control

5.5 Linked Data Graph

5.6 Discourse Analysis

Chapter 6

Evaluation

Hello

6.1 Experiment Design

6.1.1 Calculating Sample Size

6.1.2 Selecting Questions

6.2 Data Collection

6.2.1 Questionnaire Design

6.2.2 Dataset Frequency Distribution

6.3 Results

6.3.1 Keywords T-Test

6.3.2 Q&A Correlation Test

6.3.3 Summary

Chapter 7

Conclusions

It works.

7.1 Summary of Results

7.2 Limitation of System

7.3 Future Work

Appendix A

Stuff

The following gets in the way of the text....

