



**Universidade do Minho**

Escola de Engenharia

Departamento de Informática

Jorge Caldas

## **Analysis and Visualisation of Dynamic Social Networks**

September 2016



**Universidade do Minho**

Escola de Engenharia

Departamento de Informática

Jorge Caldas

## **Analysis and Visualisation of Dynamic Social Networks**

Master dissertation

Master Degree in Computer Science

Dissertation supervised by

**Pedro Rangel Henriques**

**Alda Lopes Ganarski**

September 2016

---

## ACKNOWLEDGEMENTS

---

Write acknowledgements here

---

## ABSTRACT

---

Write abstract here (en) or import corresponding file

---

## RESUMO

---

Escrever aqui resumo (pt) ou importar respectivo ficheiro

---

## CONTENTS

---

1	INTRODUCTION	1
1.1	Context and Problem	1
1.2	Motivation	1
1.3	Goals	1
2	SOCIAL NETWORKS IN SOCIOLOGY	2
2.1	Origins of Social Networks	2
2.1.1	Sociology Perspective	3
2.2	Relevant SN related terms	3
3	ONLINE SOCIAL NETWORKS	5
3.1	Portuguese and Online Social Networks	7
3.2	SN A	7
3.2.1	Domain Modeling	7
3.2.2	API	7
3.3	SN B	7
3.4	SN C	7
4	SOCIAL NETWORK ANALYSIS	8
4.1	Fundamental Concepts for Network Analysis	10
4.2	Network Analysis	11
4.2.1	Scientific Background	11
4.2.2	Power Law	11
4.2.3	Centrality Measures	11
4.2.4	Link Analysis	11
4.2.5	...	11
4.3	Six Degrees of Separation	11
4.4	Network Visualisation	11
4.5	Real World Applications	11
5	THE PROBLEM AND ITS CHALLENGES	12
6	PROPOSED SOLUTION	13
6.1	Solution Requirements	13
6.1.1	Requirements Analysis	13
6.1.2	Requirements Specification	13
6.1.3	Requirements Prioritisation	13
6.2	System Modeling	13

6.3	System Architecture	13
6.4	Technology Selection	13
6.4.1	Technology A	13
6.4.2	Technology B	13
6.4.3	Technology C	13
6.4.4	Technology Comparison	13
6.4.5	Decision	13
7	IMPLEMENTATION	14
7.1	Data Extraction	14
7.1.1	Data Sources	14
7.2	Data Mining	14
7.3	Back end	14
7.4	Front end	14
7.5	Outcomes	14
8	CASE STUDIES	15
8.1	Results	15
8.2	Discussion	15
8.3	Summary	15
9	CONCLUSION	16
9.1	Conclusions	16
9.2	Prospect for future work	16

---

## LIST OF FIGURES

---



---

## LIST OF TABLES

---

Table 1	Table of <i>Online Social Networks (OSNs)</i> ( <a href="#">statista.com</a> (2016), <a href="#">expandedramblings.com</a> (2016))	6
---------	--	---

---

## ACRONYMS

---

### O

OSN Online Social Network.

### S

SN Social Network.

SNA Social Network Analysis.

---

## INTRODUCTION

---

### 1.1 CONTEXT AND PROBLEM

### 1.2 MOTIVATION

### 1.3 GOALS

---

## SOCIAL NETWORKS IN SOCIOLOGY

---

Nowadays is hard to find something that is not organized as a network, if one tries to understand something about the world around us, then definitely one needs to know something about networks.

Curiously if you look up the term *Social Network (SN)* in the [Dictionary \(2002\)](#), we may face the following:

*"a website or computer program that allows people to communicate and share information on the Internet using a computer or mobile phone"*

But, even if today we automatically think in SNs as websites (or web applications), deep down we know when talking about SNs, we refer to a much more broader term, that said, we may consider a SNs as the following:

*"A social structure made of nodes that are generally individuals or organizations. A social network represents relationships and flows between people, groups, organizations, animals, computers or other information/knowledge processing entities. The term itself was coined in 1954 by J. A. Barnes."* [Beal \(2016\)](#)

One may say that networks work like pipes, and through them things flow, from individual to individual inside the network. It's through networks that big institutions can organize themselves, and actually add value to society despite the large number of individuals.

### 2.1 ORIGINS OF SOCIAL NETWORKS

*"The network concept is one of the defining paradigms of the modern era."* [Kilduff and Tsai \(2003\)](#)

Before talking of network from the sociology perspective, one needs to review the network concept, which is broadly used across multiple fields of study, this include, physics, biology, linguistic, anthropology, mathematics, computer science and more recently computer networks.

But why is the network approach so adopted in such diversification fields? According to Kilduff and Tsai (2003), the answer is, because networks allows us to capture the interactions of any individual unit within the larger field of activity to which the unit belongs.

### 2.1.1 Sociology Perspective

*"(...) many people attribute the first use of the term "social network" to Barnes (1954). The notion of a network of relations linking social entities, or of webs or ties among social units emanating through society, has found wide expression throughout the social sciences. (...)" Wasserman and Faust (1994)*

The SN concept has been around for many years now, maybe not in the exact format that nowadays, we are familiarized with ("web way", in a manner of speaking), but in a more abstract sense, applied in real life within real connections. Wasserman and Faust (1994), refer that this term has first came into discussion in 1954, introduced by Barnes, J.A.

*"Social relations in Bremnes, Norway, fall into three categories: relatively stable formal organizations serving many different purposes, unstable associations engaged in fishing, and interpersonal links that combine to form a social network and on which perceptions of class are based. In fishing situations, orders are given and obeyed; in the other social settings, consensus decisions are reached obliquely and tentatively." Barnes (1954)*

In the above citation, John Arundel Barnes, does a very well succeed reflection about the relationships of the people from Bremnes (Norway). The author points out that relations can form organizations for serving a specific purpose, and today we clearly see that the chosen path of SNs and also OSNs, was narrow down SNs to very specific purposes, such as professional networks. So one may say that John Arundel Barnes not only coined the term SN, but also was one of the first who described **interest-based social networks**.

## 2.2 RELEVANT SN RELATED TERMS

**In this section talk about some inherent concepts of SN, only if they are found relevant.** (Review this theories. Why are they important in sociology? What is their placement (fitting) in the thesis?)

- Homophily and Heterophily
- Structuralism
- Structural functionalism

- Conflict theories
- Social constructionism

---

## ONLINE SOCIAL NETWORKS

---

People need to connect other people, and the urge for connection, bring to us what today are known as OSNs. This web sites allows to define a profile as an individual, and to share and visualize content with other individuals in the network, therefore connecting.

*"We define Online Social Networks as web-based services that allow individuals to construct a public or semi-public profile within a bounded system, articulate a list of other users with whom they share a connection, and view and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site. Ellison et al. (2007) <sup>1</sup>*

---

<sup>1</sup> A table is presented on the next page. The blank space is due to the size of the table.

Name	Year of launch	Registered Users	Provides an API?	Description/Purpose
Facebook	2004	1 712 000 000	Yes	<b>General.</b> Photos, videos, blogs, apps.
Google+	2011	1 600 000 000	Yes	<b>General.</b> Google+ is an interest-based social network that is owned and operated by Google.
Youtube	2005	1 000 000 000	Yes	Allows billions of people to discover, watch and share originally-created videos. Provides a forum for people to connect, inform, and inspire others.
Qzone	2005	652 000 000	No	<b>General.</b> It allows users to write blogs, keep diaries, send photos, listen to music, and watch videos. It's only available in Chinese.
Twitter	2006	645 750 000	Yes	<b>General.</b> Micro-blogging, RSS, updates.
Tumblr	2007	555 000 000	Yes	Microblogging platform and social networking website.
Instagram	2010	300 000 000	Yes	A photo and video sharing site.
Sina Weibo	2009	300 000 000	Yes	Social microblogging site in mainland China.
VK	2006	249 409 900	Yes	<b>General,</b> including music upload, listening and search. Popular in Russia and former Soviet republics.
LinkedIn	2003	200 000 000	Yes	Business and professional networking.
Vine	2013	200 000 000	No	Short-form video sharing service where users can share six-second-long looping video clips.
Pinterest	2010	176 000 000	Yes	The worlds catalog of ideas. Find and save recipes, parenting hacks, style inspiration and other ideas to try.
Reddit	2005	35 000 000	Yes	Social media, social news aggregation, web content rating, and discussion website.
Flickr	2007	32 000 000	Yes	Helping people make their photos available to the people who matter to them. Enable new ways of organising photos and video.
Meetup	2002	27 590 000	Yes	World's largest network of local groups. Meetup makes it easy for anyone to organize a local group or find one of the thousands already meeting up face-to-face. <a href="http://meetup.com">meetup.com</a> (2016)
Couchsurfing	2004	12 000 000	No	Couchsurfing connects travellers with a global network of people willing to share in profound and meaningful ways, making travel a truly social experience. Is commonly used by travellers to find free hosts across the globe. <a href="http://couchsurfing.com">couchsurfing.com</a> (2016)
ResearchGate	2008	10 000 000	No	Built by scientists, for scientists. Connect the world of science and make research open to all. <a href="http://researchgate.net">researchgate.net</a> (2016)

Table 1: Table of OSNs ([statista.com](http://statista.com) (2016), [expandedramblings.com](http://expandedramblings.com) (2016))



The Table 1 lists the most used and popular OSNs, ordered by the estimated number of registered users.

The first obvious comment on the listed OSNs is that general purpose OSNs have more users (social networks with the word *General* in bold), being Youtube an exception, since it is not a general purpose OSNs, neither is focused on individuals, it is build around **social objects**, the videos.

The grey scale in the first column of Table 1 divides OSNs in three groups: the first and smallest, the 1 billion or more users OSNs; the second the OSNs with less than 1 billion users and more then 100 million; finally, the third group, OSNs with less then 100 million users. At this point, we begin to observe that **the narrower purpose OSNs** such as ResearchGate (mainly for researchers) or Couchsurfing (mainly for open minded travellers), **have a smaller number of registered users**, which is expected since the target audience is also smaller.

Other OSNs not listed in the Table 1, but still worth mentioning include **Classmates** (helps users finding classmates form kindergarten, primary school, high school etc.) known for being one of the first OSNs, since it was launched in 1995, and **Ask.fm** (allows users to interact with other users asking and answering questions (revealing identity is optional)).

### 3.1 PORTUGUESE AND ONLINE SOCIAL NETWORKS

#### 3.2 SN A

##### 3.2.1 Domain Modeling

##### 3.2.2 API

#### 3.3 SN B

#### 3.4 SN C

---

## SOCIAL NETWORK ANALYSIS

---

*Social Network Analysis (SNA)* is the study of how people are connected to each other, basically it studies a set of relations among a set of entities, these entities may be individuals, organizations, or even countries.

The common analysis procedure consists in mapping the network and then create metrics to characterize the network. Then one tries to figure what is the structure of the network and why does it have that structure. SNA is also about look at the individuals inside the network and where are those individuals located.

### 4.1 FUNDAMENTAL CONCEPTS FOR NETWORK ANALYSIS

The concepts listed below are of key importance to understand SNA. Wasserman and Faust (1994)

- *Actor* - SNA is concerned with understanding the linkages among social entities and the implications of these linkages, these social entities are described as actors. Actors are discrete individual, corporate, or collective social units.
- *Relational Tie* - Actors are linked to one another through *social ties*. The type of ties may be extensive, and it describes the nature of the connection. Some example of ties:
  - **Evaluation** of one person by another;
  - **Transference** of resources (business transactions);
  - **Association** (to social event or cause);
  - **Behavioural** interactions (communicating);
  - **Moving** between places or statuses (migration, social or physical mobility);
  - Others may be: physical connection (roads, rivers), formal relations (authority), biological relationship;

- *Dyad* - The most basic relationship that can be established is a dyad, a connection between two actors.
- *Triad* - A relation established between three actors. Many studies included breaking SNs down to small groups (triads), this allowed a more clear conclusion about the transitivity of the connections.
- *Subgroup* - It defines any subset of actors in a SN (conceptually, subgroups come after dyads and triads).
- *Group* - A finite set of actors who for conceptual, theoretical or empirical reasons are treated as a finite set of individuals in which network measurements are made.
- *Relation* - A collection of ties of a specific kind among members of a group is called a **relation** (e.g. a connection in *LinkedIn* is a relation while evaluating our connections of sending them messages are ties).
- *SN* - With the definitions of actor, group and relation, a SN consists of a finite set or sets of actors and the relation or relations defined on them. The presence of relation information is critical and defining feature of a SN.

## 4.2 NETWORK ANALYSIS

### 4.2.1 *Scientific Background*

*Graphs*

*Statistics*

...

### 4.2.2 *Power Law*

### 4.2.3 *Centrality Measures*

### 4.2.4 *Link Analysis*

### 4.2.5 ...

## 4.3 SIX DEGREES OF SEPARATION

## 4.4 NETWORK VISUALISATION

## 4.5 REAL WORLD APPLICATIONS

---

THE PROBLEM AND ITS CHALLENGES

---

---

## PROPOSED SOLUTION

---

### 6.1 SOLUTION REQUIREMENTS

#### 6.1.1 *Requirements Analysis*

#### 6.1.2 *Requirements Specification*

#### 6.1.3 *Requirements Prioritisation*

### 6.2 SYSTEM MODELING

### 6.3 SYSTEM ARCHITECTURE

### 6.4 TECHNOLOGY SELECTION

#### 6.4.1 *Technology A*

#### 6.4.2 *Technology B*

#### 6.4.3 *Technology C*

#### 6.4.4 *Technology Comparison*

#### 6.4.5 *Decision*

---

## IMPLEMENTATION

---

### 7.1 DATA EXTRACTION

#### 7.1.1 *Data Sources*

### 7.2 DATA MINING

### 7.3 BACK END

### 7.4 FRONT END

### 7.5 OUTCOMES

---

## CASE STUDIES

---

Application of main result (examples and case studies)

8.1 RESULTS

8.2 DISCUSSION

8.3 SUMMARY



---

## CONCLUSION

---

Conclusions and future work.

### 9.1 CONCLUSIONS

### 9.2 PROSPECT FOR FUTURE WORK

---

## BIBLIOGRAPHY

---

- John Arundel Barnes. *Class and committees in a Norwegian island parish*. Plenum New York, 1954.
- Vangie Beal. Webopedia definition for social network, 2016.
- couchsurfing.com. Couchsurfing about page, 2016.
- Cambridge Dictionary. Cambridge dictionaries online, 2002.
- Nicole B Ellison et al. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2007.
- expandedramblings.com. Social media statistics, 2016.
- Martin Kilduff and Wenpin Tsai. *Social networks and organizations*. Sage, 2003.
- meetup.com. Meetup about page, 2016.
- researchgate.net. Researchgate about page, 2016.
- statista.com. statista, global social media ranking, 2016.
- Stanley Wasserman and Katherine Faust. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.