



Natural Language Processing Lab, Soochow University, Suzhou, China  
 {wangzq.antony, shoushan.li}@gmail.com, {kongfang, gdzhou}@suda.edu.cn

## Introduction

- In this paper, we address the task of *personal profile summarization* by leveraging both personal profile textual information and social networks.
- we propose a *collective factor graph (CoFG) model* to incorporate all these resources of knowledge to summarize personal profiles with *local textual attribute functions* and *social connection factors* (e.g. *co-major, co-university, and co-corporation*).

## Data Collection

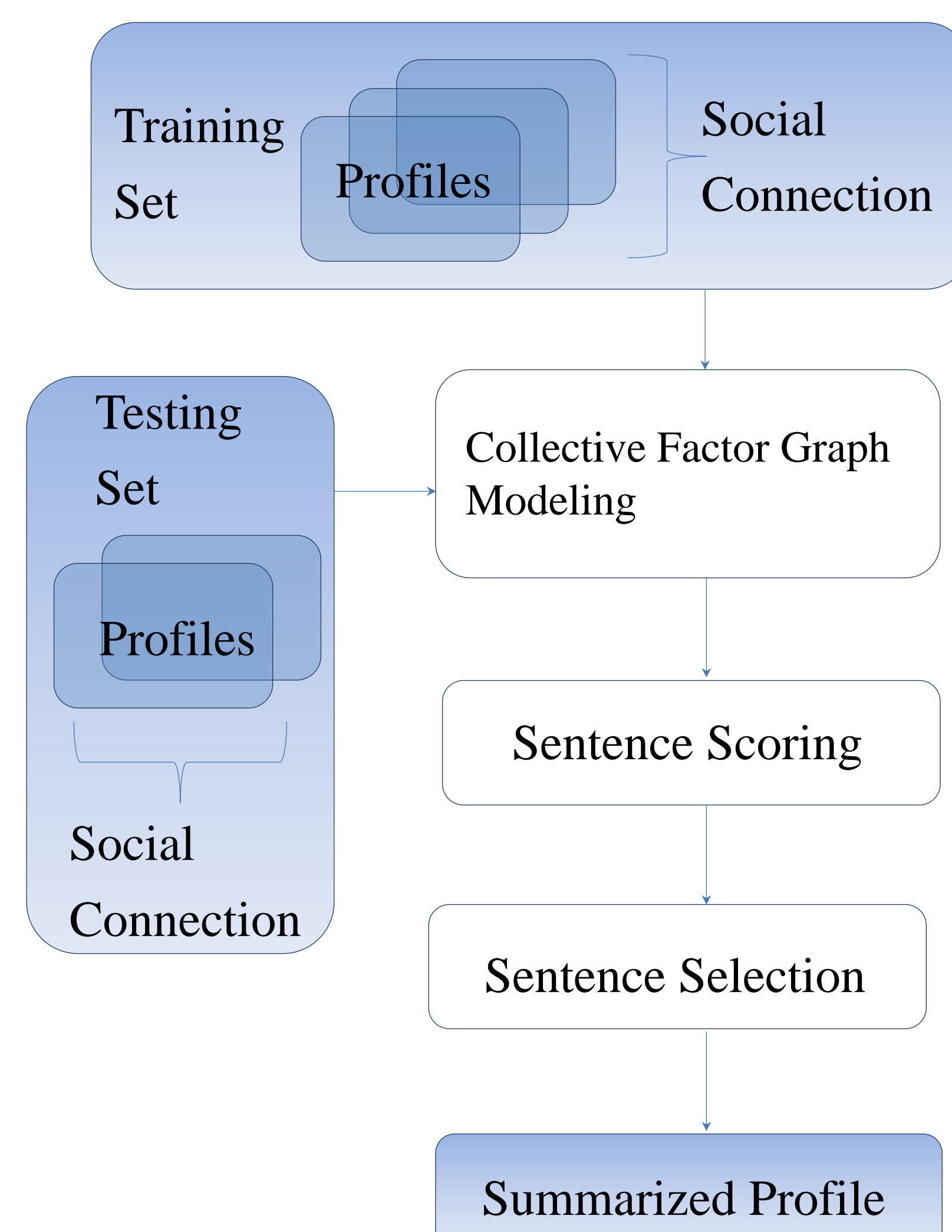
Figure: A profile page from LinkedIn.com

John Smith	
Overview	
<b>Current</b>	Applied Researcher at Apple Inc.
<b>Previous</b>	Senior Research Scientist at IBM
<b>Education</b>	MIT, Georgia Institute of Technology,
Summary	
Machine learning researcher and engineer on many fields: Query understanding. Information extraction	
Experience	
Applied Researcher Apple Inc., September 2012 ~ Query recognition and relevance	
Education	
MIT Ph.D., Electrical Engineering, 2002 – 2008	

- The text of *Summary field* in the profiles are served as the reference **summary (the standard answers)**.
- We treat the text of *Experience field* as the source of summary for each profile.
- We collect *social context information* from *Education* and *Experience* field, and these social contexts are including by LinkedIn explicitly.

## Collective Factor Graph Model

Figure: Overview of our framework



- we formalize the problem of personal profile summarization in a **supervised pairwise factor graph model**.
- We propose an approach referred to as *Loopy Belief Propagation algorithm* to learn the model for generating the summary of the profile.

## Model Definition

### Probability of graph model

$$P(Y|X, G) \propto P(Y|G) \prod_i P(x_i | y_i)$$

### Local textual attribute functions

$$P(x_i | y_i) = \frac{1}{Z_1} \exp \left\{ \sum_{j=1}^d \alpha_j f_j(x_{ij}, y_i) \right\}$$

### Social connection factor function

$$P(Y|G) = \frac{1}{Z_2} \exp \left\{ \sum_i \sum_{j \in NB(i)} g(i, j) \right\}$$

$$g(y_i, y_j) = \exp \left\{ \beta_{ij} (y_i - y_j)^2 \right\}$$

## Experimentation

- In the experiment, we use the corpus collected from LinkedIn.com that *contains 497 profiles*.
- we *extract 40 words to construct the summary* for each profile. We use 200 personal profiles as the testing data, and the remaining ones as the training data.
- We use the ROUGE-1.5.5 (Lin and Hovy, 2004) toolkit for evaluation.

Table: Performances of different approaches to profile summarization

	ROUGE-2	ROUGE-L	ROUGE-W	ROUGE-SU4
Random	0.0219	0.1363	0.0831	0.0288
HITS	0.0295	0.1499	0.0905	0.0355
PageRank	0.0307	0.1574	0.0944	0.0383
MaxEnt	0.0349	0.1659	0.0995	0.0377
<b>CoFG</b>	<b>0.0383</b>	<b>0.1696</b>	<b>0.1015</b>	<b>0.0415</b>