

APPLICATION OF SOCIAL SIGNALS IN MULTIMEDIA SOCIAL NETWORKS

SUMIT NEGI



**AMAR NATH AND SHASHI KHOSLA SCHOOL OF
INFORMATION TECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY DELHI
OCTOBER 2020**

© Indian Institute of Technology Delhi (IITD), New Delhi, 2020

APPLICATION OF SOCIAL SIGNALS IN MULTIMEDIA SOCIAL NETWORKS

by

SUMIT NEGI

Amar Nath And Shashi Khosla School of Information Technology

Submitted

in fulfilment of the requirements of the degree of Doctor of Philosophy

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

OCTOBER 2020

Certificate

This is to certify that the thesis titled **APPLICATION OF SOCIAL SIGNALS IN MULTI-MEDIA SOCIAL NETWORKS** being submitted by **Mr. SUMIT NEGI** to the Amar Nath Shashi Khosla School of Information Technology, Indian Institute of Technology Delhi, for the award of **Doctor of Philosophy** is a record of bona-fide research work carried out by him under our guidance and supervision. In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree. The work presented in this thesis has not been submitted elsewhere, either in part or full, for the award of any other degree or diploma.



Professor Santanu Chaudhury
Department of Electrical Engineering
Indian Institute of Technology Delhi
New Delhi - 110 016
(Advisor)



Professor Prem K Kalra
Department of Computer Science & Engineering
Indian Institute of Technology Delhi
New Delhi - 110 016
(co-advisor)

Acknowledgements

I would like to thank my family and my teachers for giving me the courage and encouragement to pursue this program. I would like to tender my heartfelt thankfulness to my supervisor Prof. Santanu Chaudhury for his supervision, continuous appreciation and enthusiastic support in every stage of this work. His deep insight into problems, ideas, technical guidance and encouragement was a constant source of inspiration and motivation. His critical comments on my technical writing, details on every aspect of this work has guided me towards the completion of this thesis. I am extremely grateful to the members of my Student Research Committee (SRC), Prof Niladri Chatterjee, Prof Rahul Garg, Prof Sumeet Agarwal, Prof Prem Kumar Kalra. I would like to specially thank Prof Prem Kumar Kalra for his support in this pursuit.

Finally, my wife Udipi Ghosh has been a constant support throughout this journey. Her patience, sacrifice, inspiration and support are of special mention. My parents and in-laws have showed immense patience and support during the course of my work. I dedicate my thesis to them.

Sumit Negi

Abstract

There has been an exponential growth in the amount of multimedia content that is being generated on multimedia social networks such as YouTube, Flickr, Twitter etc. Unlike traditional multimedia content, content shared on these networks is annotated with a rich set of network and social metadata features which are generated as a byproduct of content creation and sharing on these networks. This social and network metadata (which we refer to as *social signals*) includes tags, description, geo-location, subscription, community membership, informal interactions via comments etc. As part of this thesis we explore how applications such as community detection, link prediction and image search can leverage social signals to their advantage. Sparsity, ambiguity and heterogeneity of these features pose challenges which are addressed using a wide array of machine learning methods such as probabilistic modelling, multi-task/multi-view learning and multiple kernel methods. Experiments on real world data sets confirm our hypothesis that incorporating social signals improves results for different multimedia tasks.

अनुसंधान सार

मल्टीमीडिया सोशल नेटवर्क जैसे कि यूट्यूब, फ़्लिकर, ट्विटर आदि की वजह से मल्टीमीडिया सामग्री की मात्रा में एक घातीय वृद्धि हुई है। इन मल्टीमीडिया सोशल नेटवर्कों पर उत्पन्न सामग्री में समृद्ध सामाजिक मेटाडेटा विशेषताएं हैं जो सामग्री निर्माण और साझा करने के कारण उत्पन्न होती हैं। यह सामाजिक और नेटवर्क मेटाडेटा (जिसे हम सामाजिक संकेतों के रूप में संदर्भित करते हैं) हैं टैग, विवरण, भू-स्थान, सदस्यता, सामुदायिक सदस्यता आदि। इस थीसिस में हम अनुसंधान करते हैं कि इस मेटाडेटा का उपयोग विभिन्न समस्याओं जैसे कि सामुदायिक पहचान, लिंक भविष्यवाणी और छवि खोज के लिए कैसे किया जा सकता है। वास्तविक डेटा सेट पर प्रयोग हमारी परिकल्पना की पुष्टि करते हैं सामाजिक संकेतों को शामिल करने से विभिन्न मल्टीमीडिया कार्यों के लिए परिणाम बेहतर होते हैं।

Contents

List of Figures	xi
List of Tables	xiii
1 Introduction	1
1.1 Objective and Challenges	1
1.2 Applications	2
1.3 Contributions	2
1.4 Overview and Chapter Layout	4
2 Literature Review	7
2.1 Motivation	7
2.1.1 Community Discovery	9
2.1.2 Link Prediction	10
2.1.3 Social Multimedia Retrieval	10
2.2 Conclusion	11
3 Joint Community and Topic Discovery	13
3.1 Introduction	13
3.2 Community Discovery in YouTube	14
3.3 Modeling interactions and content metadata	15
3.3.1 Inference	16
3.4 Experiments: Results and Discussion	19
3.4.1 YouTube Data Set	19
3.4.2 Results	20
3.4.3 Discussion	21
3.5 Conclusion	22
4 Partial Multi-View Clustering	25
4.1 Introduction	25
4.2 Multi-View Clustering	26
4.3 Graph Regularized Partial Multi-View Clustering	27
4.3.1 Notations	27

4.3.2	GPMVC Formulation	28
4.3.3	GPMVC Updates	30
4.4	Experiment	31
4.4.1	Partial View Experiments	31
4.5	Conclusion	36
5	Link Prediction in Multimedia Social Networks	37
5.1	Introduction	37
5.2	Link Prediction	38
5.2.1	Link Prediction in Heterogenous Networks	39
5.2.2	Metric Learning	39
5.2.3	Our Contributions	40
5.3	Definition and Notations	41
5.4	Structure Preserving Metric Learning	42
5.5	Robust Multi-Task SPML	43
5.5.1	Robust Metric Learning	43
5.5.2	Degree Distributional Metric Learning	43
5.5.3	Learning Task Correlation	44
5.6	Optimization	45
5.6.1	Link Prediction	47
5.7	Experiments	48
5.7.1	Baseline	48
5.7.2	Evaluation Metric	49
5.7.3	Dataset and Experiment Setup	49
5.7.4	Results	51
5.8	Conclusion	52
6	Search Result Diversification	53
6.1	Introduction	53
6.2	Image Search Result Diversification	54
6.2.1	Main Contributions	55
6.3	System Description	56
6.3.1	Topic Richness	56
6.3.2	MVC: multi-view clustering	57
6.3.3	UPD: User profile based diversity	58
6.3.4	Flickr baseline	59
6.4	Data Set Description	59
6.5	Experiment	60
6.5.1	Evaluation Measure	60
6.5.2	Evaluation Measure and Results	60
6.5.3	Experiment Setup	61
6.5.4	Observations	62
6.6	Conclusion	62

7	Identifying Diverse Images in Flickr	63
7.1	Introduction	63
7.2	Visual Diversification	64
7.3	Problem Formulation and Notations	65
7.4	Candidate Similarity Measure	65
7.4.1	Visual Feature based Similarity	65
7.4.2	Flickr Group based Similarity	66
7.4.3	Tag based Similarity	67
7.5	Metric Learning with Multiple Kernel	67
7.6	Constraint Generation	69
7.7	Identifying Diverse Images	71
7.8	Experiment	71
7.8.1	Data Preparation	71
7.8.2	Similarity calculation	72
7.8.3	Evaluation Measure	73
7.8.4	Observation	73
7.9	Conclusion	74
8	Conclusion	75
8.1	Contributions	75
8.2	Future Work	76
A	Appendix	77
	Appendices	77
	Bibliography	81
	Publications	89
	Biography	91

List of Figures

1.1	Tasks that can benefit from analysis of social signals	3
3.1	Jointly modeling interaction and content metadata	16
3.2	Symbols	17
3.3	YouTube user communities and topics (top row <i>Howto & DIY</i> , bottom row <i>News</i>)	21
3.4	Perplexity results (best viewed in color)	22
4.1	View-to-Instance Mapping (P) and Inverse Mapping (P^-)	28
4.2	NMI VALUES VS. PER FOR TWO VIEW DATASETS (BEST VIEWED IN COLOR)	34
4.3	NMI VALUES VS. PER FOR MULTI-VIEW DATASETS (BEST VIEWED IN COLOR)	36
5.1	Decompose a heterogeneous network into homogeneous networks (where each network captures interactions of a single type) and then jointly learn <i>link-type distance measures</i> for each network. The learnt metrics are then used for the corresponding link prediction task. For example the learnt metric d_{G1} is used to infer which two (existing or new) users are likely to form a link.	39
6.1	Images with caption “ <i>scorpion</i> ”	54
7.1	Approach Overview	64
7.2	Label Generation	70
7.3	CR@10 for 10 queries	74

List of Tables

3.1	List of social signals collected in the YouTube data set	19
3.2	YouTube Dataset <i>Howto & DIY</i> and <i>News</i>	20
4.1	Details of the datasets	31
4.2	NMI numbers on ORL (2 view)	34
4.3	NMI numbers on Digit (2 view)	34
5.1	Baseline method AUC metrics for different link-types – Flickr (top-table), DBLP (bottom-table) dataset	47
5.2	Proposed method AUC metrics for different link-types – Flickr (top-table), DBLP (bottom-table) dataset	49
6.1	Query Statistics	60
6.2	CR@10 Results	61
7.1	Keyword queries and possible interpretations	72