Project Topics

Below is a list of possible project topics. Some of these are open-ended, meaning that you are required to come up with a new algorithm or model, and formulate it yourselves. Such projects may require more effort, but they will be also graded based on the effort, as well as the final result. Others are more straight-forward, you would need to obtain a complex dataset and apply algorithms on this data. There are also more theoretical projects, and more practical ones, so you can pick depending on your preference.

You will also have to present in class one paper related with you project. The list below includes the paper for each project.

Papers also vary in difficulty and scope. For experimental papers, that just report results of experimental studies, we expect that you just present and explain the main findings. Since such papers require less effort, you will be asked to present 2 such papers.

Projects should be done **in teams of at most two** students.

Deliverables and Timeline:

- A web page with all information related to your project (first version: week before Christmas, final version: end of January)
- A two-page project proposal outlining what you plan to do. This should include the topic of your presentation (first version: week before Christmas)
- A 20' presentation of 1-2 research papers related to your project (presentations are scheduled for 13/1 and 20/1)
- The source code of your project (end of January)
- A final report describing your project following a specific format (end of January)

Topic 1

Ego network characterization

Project:

Collect your ego network on Facebook. We need a network with at least 100 users, so do not take this project unless you are active in Facebook and have a large enough set of friends. Using this data, perform an analysis on your ego network, such as finding communities (based on topics or profile information, e,g., location), understanding homophily, activity patterns, etc.

Paper:

Lars Backstrom, Eytan Bakshy, Jon M. Kleinberg, Thomas M. Lento, Itamar Rosenn:
Center of Attention: How Facebook Users Allocate Attention across Friends. ICWSM 2011
AND
Johan Ugander, Brian Karrer, Lars Backstrom, Cameron Marlow:
The Anatomy of the Facebook Social Graph. CoRR abs/1111.4503 (2011)

Topic 2

Local and global event identification

Project:

Use the Instagram API to collect datasets from Instagram and perform an analysis on the collected data. A possible project: Identify posts that refer to an even that is localized both in time (something that happened within a specific time interval) and in space (something that happened in a restricted area – e.g. in Ioannina).

Paper:

Tye Rattenbury, Nathaniel Good, Mor Naaman: Towards automatic extraction of event and place semantics from flickr tags. SIGIR 2007: 103-110

The following papers may be useful for your project:
Fotis Psallidas, Hila Becker, Mor Naaman, Luis Gravano: Effective Event Identification in Social Media. IEEE Data Eng. Bull. 36(3): 42-50 (2013)
Hila Becker, Mor Naaman, Luis Gravano: Beyond Trending Topics: Real-World Event Identification on Twitter. ICWSM 2011
Jon M. Kleinberg: Bursty and hierarchical structure in streams. KDD 2002: 91-101

Topic 3

Team formation

Project:

Extend the model and algorithms for team formation described in class to the case of negative edges. Implement and test them.

Paper:

Mehdi Kargar, Aijun An: Discovering top-k teams of experts with/without a leader in social networks. CIKM 2011: 985-994

Topic 4

Recommending tags for influence maximization

Project:

Often tags (or, keywords) are associated with resources (such as articles, photos, or posts). A basic criterion for selecting tags is their relevance to the resource. In this project, you will investigate another aspect. Given a set of m tags, an estimation of their relevance to the resource and the corresponding tag diffusion network, select the $k < m$ tags that would maximize the diffusion of the resource.

Paper:

Su Mon Kywe, Tuan-Anh Hoang, Ee-Peng Lim, Feida Zhu: On Recommending Hashtags in Twitter Networks. SocInfo 2012: 337-350
AND
Eva Zangerle, Wolfgang Gassler, Gunther Specht, Recommending #-Tags in Twitter

The following papers may also be useful:

Vineet Chaoji, Sayan Ranu, Rajeev Rastogi, and Rushi Bhatt. Recommendations to boost content spread in social networks. In WWW, 2012.

Paul Heymann, Daniel Ramage, Hector Garcia-Molina: Social tag prediction. SIGIR 2008: 531-538

Topic 5

Communities over time

Project:
In class we described definitions and algorithms for finding communities and dense subgraphs in networks. Consider a network that changes over time. How are communities defined in this case? The goal of this project is to define the notion of communities or density of sub-graphs over time, propose algorithms for finding such structures and study experimentally.

Paper:
Jimeng Sun, Christos Faloutsos, Spiros Papadimitriou, Philip S. Yu:
GraphScope: parameter-free mining of large time-evolving graphs. KDD 2007: 687-69

This paper may also be useful
Ravi Kumar, Jasmine Novak, Prabhakar Raghavan, Andrew Tomkins: On the bursty evolution of blogspace. WWW 2003: 568-576

Topic 6

Dense subgraphs over time

Project:
An extension of community search involves discovering communities around a specific set of nodes which are given as input. Consider extending this problem in the case of evolving graphs.

Paper:
Mauro Sozio, Aristides Gionis: The community-search problem and how to plan a successful cocktail party. KDD 2010: 939-948

This paper may also be useful:
Albert Angel, Nick Koudas, Nikos Sarkas, Divesh Srivastava: Dense Subgraph Maintenance under Streaming Edge Weight Updates for Real-time Story Identification. PVLDB 5(6): 574-585 (2012)

Topic 7

Virus propagation in evolving networks

Project:

Consider a virus propagation in a network, and assume that we want to block it by immunizing some nodes. Now what happens is the network is evolving over time. We need to decide not only who to immunize but also when to immunize them. Study this problem, propose algorithms, and perform experiments.

Paper:

Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne M. VanBriesen, Natalie S. Glance: Cost-effective outbreak detection in networks. KDD 2007: 420-42

This paper may also be useful:
Hanghang Tong, B. Aditya Prakash, Tina Eliassi-Rad, Michalis Faloutsos, Christos Faloutsos: Gelling, and melting, large graphs by edge manipulation. CIKM 2012: 245-254

Topic 8

Diversifying opinions

Project:

Consider network where the nodes have opinions, and define a notion of diversity (mixing of opinions for the network or an individual) or polarization (separation of opinions). Given an opinion formation model, and a budget on the interventions we can perform, find the best interventions to minimize polarization (or maximize diversity).

Paper:

Aristides Gionis, Evimaria Terzi, Panayiotis Tsaparas:
Opinion Maximization in Social Networks. SDM 2013: 387-395

This paper may also be useful:
Pedro Henrique Calais Guerra, Wagner Meira Jr., Claire Cardie, Robert Kleinberg:
A Measure of Polarization on Social Media Networks Based on Community Boundaries.
ICWSM 2013

Topic 9

Link recommendation with negative links

Project:

Propose algorithms and techniques for link recommendations in the presence of negative connections. The goal is not only predict the links but also their sign.

Paper:

J. Leskovec, D. Huttenlocher, J. Kleinberg. Predicting Positive and Negative Links in Online Social Networks ACM WWW International conference on World Wide Web (WWW), 2010.

Topic 10

Another option is to suggest a project of your own, based on what you have seen in the class so far, questions you may have thought of, and things that are related to your research area. In this case you should create a project proposal (initially just a paragraph or an idea) and contact us to discuss it.