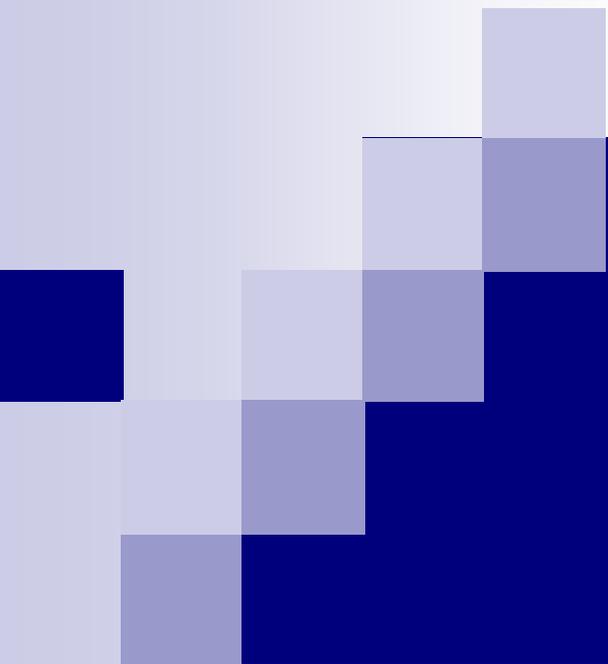


Aug 27 2010
ERC Workshop



Complex Network Modeling and Analysis

Kyomin Jung
Applied Algorithm Lab
KAIST

Applied Algorithm Lab.

■ 정교민

- Assistant Prof. in KAIST CS dept. Joint appointments in KAIST Math and EE dept.
- <http://web.kaist.ac.kr/~kyomin/>
- twitter.com/kyominj

■ Master Students

- 김보영 (Decentralized Ranking Learning)
- 임용섭 (Machine Vision, Decentralized Sensor Deportation)
- 곽남주 (Hidden Markov Model for Complex Networks)
- 허우람 (Twitter Data Analysis)
- 이슬기 (Threshold in Information Spreading)

■ Undergraduate Research Program

- 진태진 (Fast Sorting Algorithm for Partially Sorted Lists)

■ Lab Homepage: <http://aa.kaist.ac.kr/>

■ Lab Blog: <http://aalog.tistory.com/>

■ Winter School on Algorithm and Combinatorics 2010

Publications this year

Journals

- [Kyomin Jung](#), Devavrat Shah and Jinwoo Shin, *Minimizing the Rate of Convergence for Iterative Algorithms*, To appear in *IEEE Transactions on Information Theory*.
- Sung-soon Choi, [Kyomin Jung](#) and Jeong Han Kim, *Almost Tight Upper Bound for Finding Fourier Coefficients of Bounded Pseudo-Boolean Functions*, To appear in *Journal of Computer and System Sciences (JCSS)*.
- Elena Grigorescu, [Kyomin Jung](#) and Ronitt Rubinfeld, *A Local Decision Test for Sparse Polynomials*, To appear in *Information Processing Letters*.

Conference Proceedings

- Yongsub Lim, [Kyomin Jung](#) and Pushmeet Kohli, *Energy Minimization Under Constraints on Label Counts*, European Conference on Computer Vision (ECCV), September 2010, Crete, Greece.
- Arnab Bhattacharyya, Elena Grigorescu, Madhav Jha, [Kyomin Jung](#), Sofya Raskhodnikova and David Woodruff, *Lower Bounds for Local Monotonicity Reconstruction from Transitive-Closure Spanners*, International Workshop on Randomization and Computation (RANDOM) September 2010, Barcelona, Spain.
- SeongHun Lee, Minsoo Jo, [Kyomin Jung](#), and Jin Hyung Kim, *Scene Text Extraction with Edge Constraint and Text Collinearity*, International Conference on Pattern Recognition (ICPR), August 2010, Istanbul, Turkey.
- Yoram Bachrach, Reshef Meir, [Kyomin Jung](#), and Pushmeet Kohli, *Optimal Coalition Structures In Skill Games*, proceedings of AAI-10, July 2010, Atlanta, USA.

Structure Learning Algorithms

- **Learning optimal coalition structures in skill games**
 - Proceedings of AAI-10, with Yoram Bachrach, Reshef Meir, and Pushmeet Kohli
- **Energy minimization with label counting constraint**
 - ECCV 2010, with Yongsub Lim and Pushmeet Kohli
- **Scene text extraction in image**
 - ICPR 2010, with SeongHun Lee, Minsoo Jo and Jin Hyung Kim
- **Learning spanner structure in graphs**
 - Random 2010, with Arnab Bhattacharyya, Elena Grigorescu, Madhav Jha, Sofya Raskhodnikova and David Woodruff
 - SODA 2009, with Arnab Bhattacharyya, Elena Grigorescu, Sofya Raskhodnikova and David Woodruff

Algorithms for Boolean and Pseudo-Boolean function

■ Learning Pseudo-Boolean functions

- COLT(The Annual Conference on Learning Theory) 2008 , To appear in JCSS, with Sung-soon Choi and Jeong Han Kim
- To appear in IEEE Transactions on Evolutionary Computation, with Sung-soon Choi and Byung-Ro Moon

■ Threshold analysis in random Boolean formula

- Artificial Intelligence 2008, with Sung-soon Choi and Jeong Han Kim

Statistical Inference

- **Local update inference algorithm for global optimization**
 - NIPS 2009, with Pushmeet Kohli and Devavrat Shah
- **Approximate inference algorithm for minor-excluded graphs in Markov Random Field**
 - NIPS (Neural Information Processing Systems) 2007, with Devavrat Shah

Network Algorithms

- **Computing capacity region of networks**
 - INFOCOM 2008, and INFOCOM 2009, with Ramakrishna Gummadi, Devavrat Shah and Ramavarapu Sreenivas
- **Analysis of Max-weight routing & scheduling algorithm**
 - STOC 2007, with Matthew Andrews and Alexander Stolyar
- **Computing loss prob. in stochastic loss networks**
 - SIGMETRICS 2008, with Yingdong Lu, Devavrat Shah, Mayank Sharma and Mark S. Squillante

Learning Pseudo-Boolean function

$$x \in \{0,1\}^n, f(x) = \sum_{i=1}^m f_i(x_{i_1}, x_{i_2} \dots x_{i_k}) \in R$$

- Our goal is to learn f by performing function queries.
- A function query corresponds to an oracle call.
- We assume that f is a degree k Pseudo-Boolean function.
 - I.e. each f_i depends on at most k many variables.
- We design an algorithm to learn f with $O(m \log n)$ queries.

Theorem (Choi, Jung, Kim)*

$$f(x) = \sum_{i=1}^m f_i(x_{i1}, x_{i2} \dots x_{ik})$$

- We propose an adaptive, randomized algorithm that learns f with $O(m \log n)$ queries with failure probability $O(\frac{1}{n^{100}})$.

* In COLT '08, To appear in JCSS

Relation with Fourier transform

$$f(x) = \sum_{i=1}^m f_i(x_{i1}, x_{i2} \dots x_{ik})$$

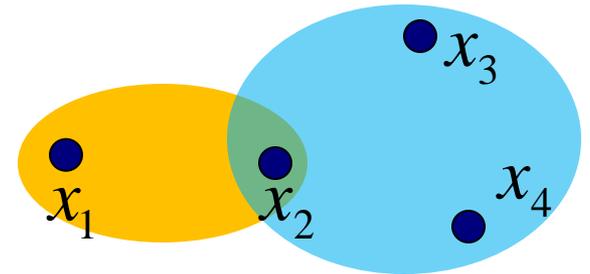
- f can be expressed by the **Fourier expression**:

$$f(x) = \sum_{H \subset [n]} \hat{f}(H) \cdot \Psi_H(x),$$

$$\Psi_H(x) = (-1)^{\sum_{i \in H} x_i}.$$

Our approach

$$f(x) = x_1x_2 - 2x_2x_3x_4$$

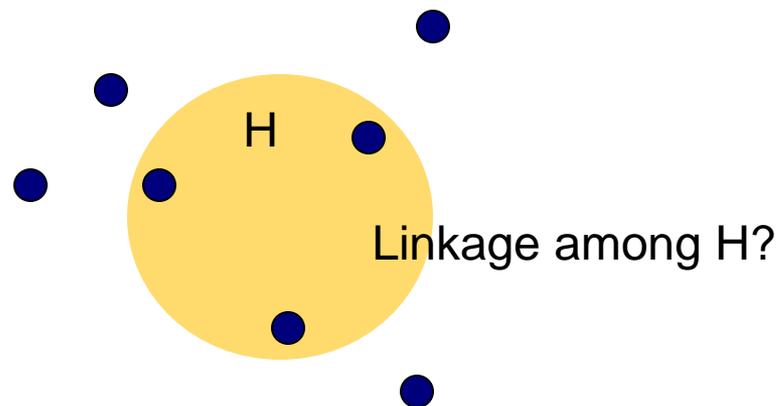


- $H \subset [n]$ is said to have **linkage**
 - If there is correlation among the variables of H
 - The hyper-graph consisting of all such H 's is called **linkage graph** of f
- First we **learn the linkage graph** of f
- Then **compute all the non-zero Fourier coefficients** of f

Learning Linkage Graph

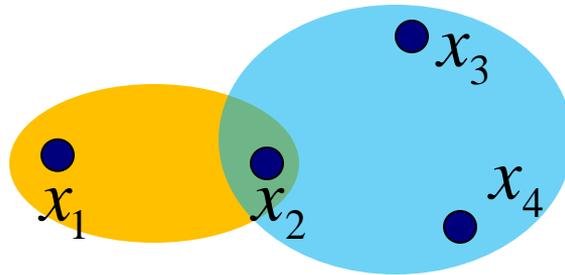
- Heckendorn and Wright [’03] has proposed the following linkage test function which tests whether there is a linkage among $H \subset [n]$.

$$L(f, H, x) = \sum_{A \subset H} (-1)^{|A|} f(x \oplus 1_A)$$



Learning Linkage Graph

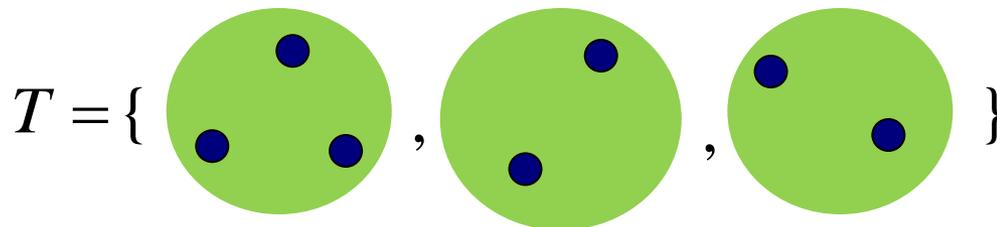
- Our key ideas are
 - Utilizing hierarchical properties of the linkage graph.



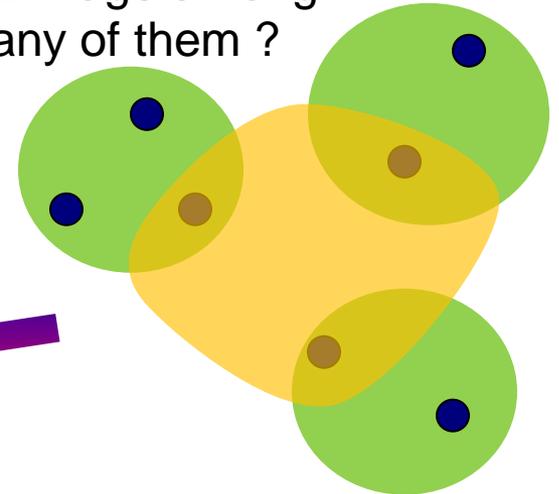
Learning Linkage Graph

- Our key ideas are
 - Utilizing hierarchical properties of the linkage graph.
 - Devising a linkage test function for group of vertices.

$$L^*(f, T, x) = \sum_{H \subset T} (-1)^{|H|} f(x \oplus (\bigoplus_{A \in H} 1_A))$$



Linkage among any of them ?



Future Research Plan

- Design and analysis of algorithms for complex networks by understanding structural properties.
- Application of machine learning to large scale complex network and computer vision, software verification etc.
- Understanding information spreading phenomenon.
- Inference in human networks : Cyworld, Twitter etc.

Thank you