Optimal Solutions for Hard Network Problems in Bioinformatics

Falk Hüffner

joint work with

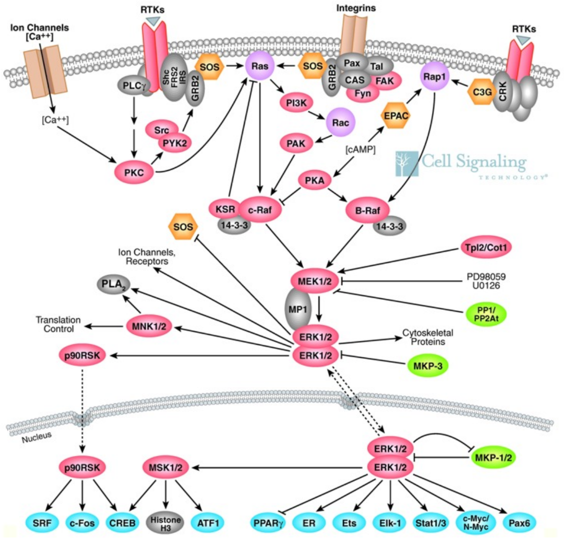
Nadja Betzler Hannes Moser Christian Komusiewicz
Rolf Niedermeier Sebastian Wernicke Thomas Zichner

Friedrich-Schiller-Universität Jena
Institut für Informatik

Computational Genomics Research Seminar
Tel Aviv University
31 October 2007

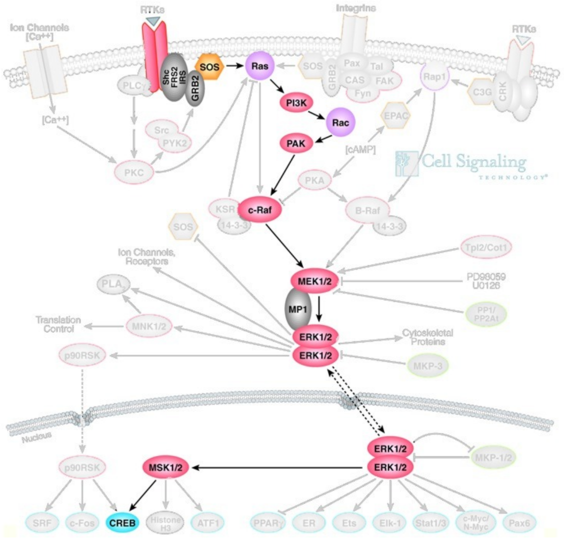
Outline

Signaling pathways



[www.cellsignal.com]

Signaling pathways



[www.cellsignal.com]

Signaling pathways

Definition

A signaling pathway is a cascade of successive protein interactions that the cell uses to react to stimuli.

Definition

A linear signaling pathway contains each protein only once.

Signaling pathways

Definition

A signaling pathway is a cascade of successive protein interactions that the cell uses to react to stimuli.

Definition

A linear signaling pathway contains each protein only once.

Linear signaling pathways

- are easy to understand and analyze;
- can serve as a seed structure for experimental investigation of more complex mechanisms.

Signaling pathways

Definition

A signaling pathway is a cascade of successive protein interactions that the cell uses to react to stimuli.

Definition

A linear signaling pathway contains each protein only once.

Linear signaling pathways

- are easy to understand and analyze;
- can serve as a seed structure for experimental investigation of more complex mechanisms.

Goal

Automated discovery of linear signaling pathways

[Steffen et al., BMC Bioinf. 2002]

Graph model

Protein interaction network:

- Proteins are nodes
- Interactions are undirected edges
- Edges are annotated with interaction probability (obtained e. g. by two-hybrid screening)

Graph model

Protein interaction network:

- Proteins are nodes
- Interactions are undirected edges
- Edges are annotated with interaction probability (obtained e. g. by two-hybrid screening)

Definition

A linear signaling pathway is a sequence of distinct proteins, where each interacts strongly with the previous one.

Graph model

Protein interaction network:

- Proteins are nodes
- Interactions are undirected edges
- Edges are annotated with interaction probability (obtained e. g. by two-hybrid screening)

Definition

A linear signaling pathway is a sequence of distinct proteins, where each interacts strongly with the previous one.

MOST PROBABLE PATH [SCOTT et al., J. Comp. Biol. 2006]

Input: Graph $G = (V, E)$, interaction probabilities $p : E \to [0, 1]$, integer $k > 0$.
Task: Find a non-overlapping path $v_1, \ldots, v_k$ of length k in G that maximizes $p(v_1, v_2) \cdot \ldots \cdot p(v_{k-1}, v_k)$.

Graph model

MOST PROBABLE PATH [SCOTT et al., J. Comp. Biol. 2006]

Input: Graph $G = (V, E)$, interaction probabilities $p : E \to [0, 1]$, integer $k > 0$.

Task: Find a non-overlapping path $v_1, \ldots, v_k$ of length k in G that maximizes $p(v_1, v_2) \cdot \ldots \cdot p(v_{k-1}, v_k)$.

Graph model

Most Probable Path [Scott et al., J. Comp. Biol. 2006]

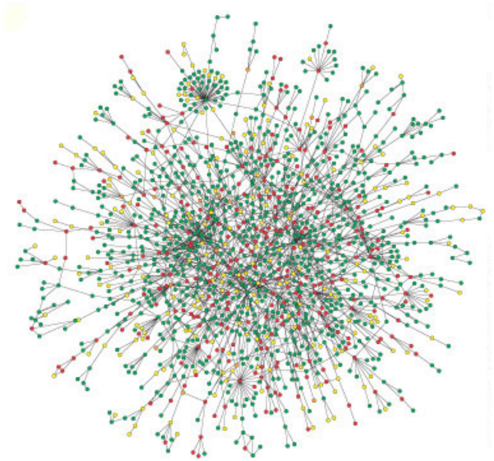
Input: Graph $G = (V, E)$, interaction probabilities $p : E \to [0, 1]$, integer $k > 0$.
Task: Find a non-overlapping path $v_1, \ldots, v_k$ of length k in G that maximizes $p(v_1, v_2) \cdot \ldots \cdot p(v_{k-1}, v_k)$.

Setting $w(e) := -\log(p(e))$:

Minimum-Weight Path

Input: Graph $G = (V, E)$, weights $w : E \to \mathbb{R}_+$, integer $k > 0$.
Task: Find a non-overlapping path $v_1, \ldots, v_k$ of length k in G that minimizes $w(v_1, v_2) + \cdots + w(v_{k-1}, v_k)$.

Example: yeast network



4 400 proteins, 14 300 interactions, looking for paths of length 5–15

Minimum-Weight Path

Theorem

MINIMUM-WEIGHT PATH *is NP-hard* [GAREY & JOHNSON 1979].

Minimum-Weight Path

Theorem

MINIMUM-WEIGHT PATH *is NP-hard* [GAREY & JOHNSON 1979].

Idea

Exploit the fact that the paths sought for are rather short ($\approx$ 5–15): restrict the exponential part of the runtime to k (parameterized complexity).

Fixed-parameter tractability

Parameterized complexity is an approach to finding exact solutions to NP-hard problems by confining the combinatorial explosion to a parameter.

Fixed-parameter tractability

Parameterized complexity is an approach to finding exact solutions to NP-hard problems by confining the combinatorial explosion to a parameter.

Definition

A problem is called fixed-parameter tractable with respect to a parameter k if an instance of size n can be solved in $f(k) \cdot n^{O(1)}$ time for an arbitrary function f.

Color-coding

Color-coding is a method for solving MINIMUM-WEIGHT PATH that is fast for short path lengths.

Color-coding [ALON, YUSTER & ZWICK, J. ACM 1995]

- randomly color each vertex of the graph with one of k colors
- hope that all vertices in the subgraph searched for obtain different colors (colorful)
- solve the MINIMUM-WEIGHT PATH under this assumption (which is much quicker)
- repeat these trial until it is reasonably certain that the path was colorful at least once

Result: exponential part of the runtime depends only on k

Dynamic programming for Minimum-Weight Colorful Path

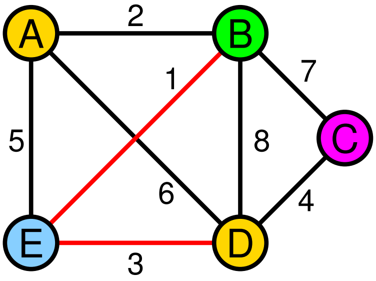
Idea

Table entry $W[v, C]$ stores the minimum-weight path that ends in v and uses exactly the colors in C.

Dynamic programming for Minimum-Weight Colorful Path

Idea

Table entry $W[v, C]$ stores the minimum-weight path that ends in v and uses exactly the colors in C.



$W[B, \{$🔵, 🟢, 🟡$\}] = 4$

Dynamic programming for Minimum-Weight Colorful Path

Coloring $c : V \to \{1, \ldots, k\}$

Recurrence

$$W[v, C] = \min_{u \in N(v)} (W[u, C \setminus \{c(v)\}] + w(u, v))$$

Dynamic programming for Minimum-Weight Colorful Path

Coloring $c : V \to \{1, \ldots, k\}$

Recurrence

$$W[v, C] = \min_{u \in N(v)} (W[u, C \setminus \{c(v)\}] + w(u, v))$$

- Each table entry can be calculated in $O(n)$ time
- $n \cdot 2^k$ table entries

$\rightsquigarrow$ Running time per trial: $O(2^k \cdot n^2)$

Color-coding running time

- $O(2^k \cdot n^2)$ time per trial
- To obtain error probability ε, one needs $O(-\ln \varepsilon \cdot e^k)$ trials

Theorem ([Alon et al., JACM 1995])

Minimum-Weight Path *can be solved in* $O(-\ln \varepsilon \cdot 5.44^k |G|)$ *time.*

Implementations of color-coding

- Find minimum-weight paths of length 10 in the yeast protein interaction networks within 3 hours ($n = 4\,400$, $k = 10$)

  [Scott et al., J. Comp. Biol. 2006]
- Pathway queries

  [Shlomi et al., BMC Bioinformatics 2006]
- Tree queries

  [Dost et al., RECOMB 2007]
- Protein docking

  [Mayrose et al., Nucleic Acids Research 2007]
- Balanced paths

  [Cappanera & Scutellà, INOC 2007]
- Automated text headline generation

  [Deshpande et al., NAACL HLT 2007]

Increasing the number of colors

Idea

Use $k + x$ colors instead of k colors.

Trial runtime:

$$O(2^k|G|) \to O(2^{k+x}|G|)$$

Increasing the number of colors

Idea

Use $k + x$ colors instead of k colors.

Trial runtime:

$$O(2^k|G|) \to O(2^{k+x}|G|)$$

Probability P_c for colorful path ($k = 8$, $\varepsilon = 0.001$):

x	0	1	2	3	4	5
P_c	0.0024	0.0084	0.0181	0.0310	0.0464	0.0636
trials	2871	816	378	220	146	106

Increasing the number of colors

Idea

Use $k + x$ colors instead of k colors.

Trial runtime:

$$O(2^k|G|) \rightarrow O(2^{k+x}|G|)$$

Probability P_c for colorful path ($k = 8$, $\varepsilon = 0.001$):

x	0	1	2	3	4	5
P_c	0.0024	0.0084	0.0181	0.0310	0.0464	0.0636
trials	2871	816	378	220	146	106

Theorem

MINIMUM-WEIGHT PATH *can be solved in* $O(-\ln \varepsilon \cdot 4.32^k|G|)$ *time by choosing* $x = 0.3k$.

Increasing the number of colors

Idea

Use $k + x$ colors instead of k colors.

Trial runtime:

$$O(2^k|G|) \to O(2^{k+x}|G|)$$

Probability P_c for colorful path ($k = 8$, $\varepsilon = 0.001$):

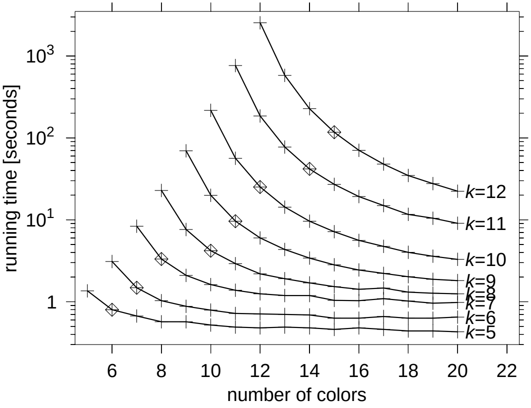
x	0	1	2	3	4	5
P_c	0.0024	0.0084	0.0181	0.0310	0.0464	0.0636
trials	2871	816	378	220	146	106

Theorem

MINIMUM-WEIGHT PATH *can be solved in* $O(-\ln\varepsilon \cdot 4.32^k|G|)$ *time by choosing* $x = 0.3k$.

But: Higher memory usage

Increasing the number of colors



Running times for the yeast protein interaction network (highlighted point of each curve marks worst-case optimum)

State space search

Idea

Consider the dynamic programming as a state space search problem (shortest path problem in an implicitly defined graph).

From a state (u, C), we can go to $(v, C \cup \{c(v)\})$ for $v \in N(u)$ with $c(v) \neq c(u)$, at cost $w(u, v)$.

State space search

Idea

Consider the dynamic programming as a state space search problem (shortest path problem in an implicitly defined graph).

From a state (u, C), we can go to $(v, C \cup \{c(v)\})$ for $v \in N(u)$ with $c(v) \neq c(u)$, at cost $w(u, v)$.

Lower bounds

- can be used to prune states
- can guide the search (A*)

State space search

Idea

Consider the dynamic programming as a state space search problem (shortest path problem in an implicitly defined graph).

From a state (u, C), we can go to $(v, C \cup \{c(v)\})$ for $v \in N(u)$ with $c(v) \neq c(u)$, at cost $w(u, v)$.

Lower bounds

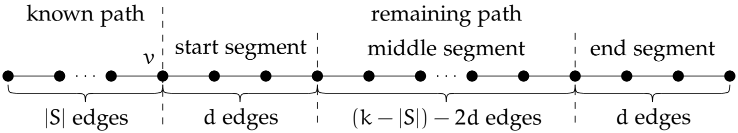
- can be used to prune states
- can guide the search (A*)

Simple lower bound:

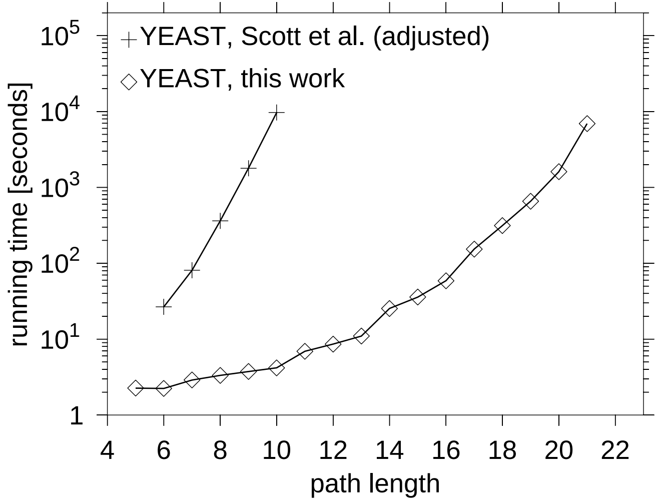
$$\text{weight} + (\text{minimum edge weight} \cdot \text{edges left})$$

Precalculated lower bounds

For each vertex u and a range of lengths $1 \leq i \leq d$, determine the minimum weight of a path of i edges that starts at u.



Yeast network

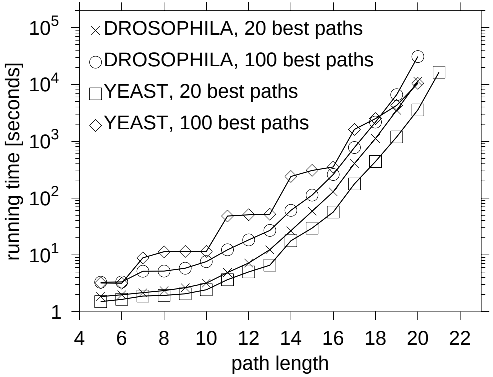


Network Comparison

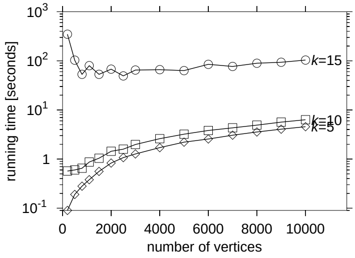
	$\lvert V\rvert$	$\lvert E\rvert$	clust. coeff.	avg. degree	max. degree
	4 389	14 319	0.067	6.5	237
	7 009	20 440	0.030	5.8	175

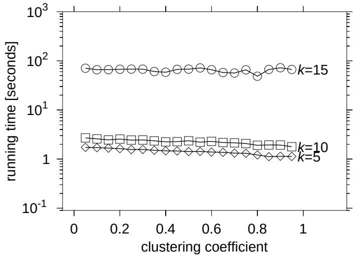
Network Comparison

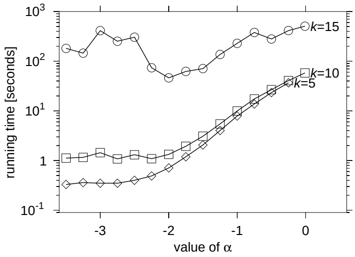
	$\lvert V\rvert$	$\lvert E\rvert$	clust. coeff.	avg. degree	max. degree
	4 389	14 319	0.067	6.5	237
	7 009	20 440	0.030	5.8	175

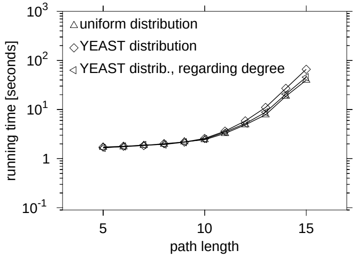


Simulations: Robustness of Algorithm







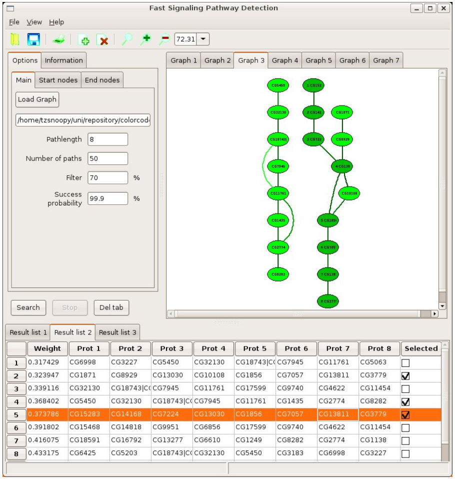


Pathway Query

Queries of *S. cerevisiae* pathways in the *D. melanogaster* network

Path length	Avg. Time [s]	Max. Time [s]	Successful Queries
4	2.24	2.57	98%
5	2.33	3.61	93%
6	3.00	23.02	81%
7	4.52	93.32	52%
8	7.49	225.61	31%
9	11.38	245.78	13%

Graphical user interface: FASPAD


Fast Signaling Pathway Detection
File View Help
72.31
Options Information
Main Start nodes End nodes
Load Graph
Graph 1 Graph 2 Graph 3 Graph 4 Graph 5 Graph 6 Graph 7
Search Stop Del tab
Result list 1 Result list 2 Result list 3

/home/tzsnoopy/uni/repository/colorcod

Pathlength 8

Number of paths 50

Filter 70 %

Success probability 99.9 %

	Weight	Prot 1	Prot 2	Prot 3	Prot 4	Prot 5	Prot 6	Prot 7	Prot 8	Selected
1	0.317429	CG6998	CG3227	CG5450	CG32130	CG18743|C(	CG7945	CG11761	CG5063	☐
2	0.323947	CG1871	CG8929	CG13030	CG10108	CG1856	CG7057	CG13811	CG3779	✓
3	0.339116	CG32130	CG18743|C(	CG7945	CG11761	CG17599	CG9740	CG4622	CG11454	☐
4	0.368402	CG5450	CG32130	CG18743|C(	CG7945	CG11761	CG1435	CG2774	CG8282	✓
5	0.373786	CG15283	CG14168	CG7224	CG13030	CG1856	CG7057	CG13811	CG3779	✓
6	0.391802	CG15468	CG14818	CG9951	CG6856	CG17599	CG9740	CG4622	CG11454	☐
7	0.416075	CG18591	CG16792	CG13277	CG6610	CG1249	CG8282	CG2774	CG1138	☐
8	0.433175	CG6425	CG5203	CG18743|C(	CG32130	CG5450	CG3183	CG6998	CG3227	☐

Free software, available at
http://theinf1.informatik.uni-jena.de/faspad/

Conclusion & Outlook

Color-coding, with some algorithm engineering, is a practical method for finding signaling pathways in protein interaction networks.

Conclusion & Outlook

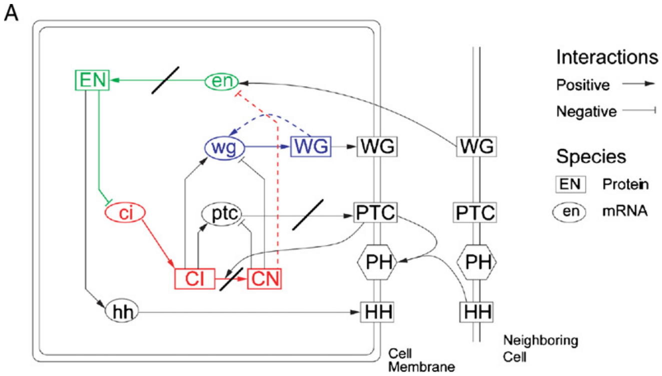
Color-coding, with some algorithm engineering, is a practical method for finding signaling pathways in protein interaction networks.

Future work:

- Richer motifs (cycles, trees, . . .)
- "Divide-and-color" [KNEIS et al., WG 2007; Chen et al., SODA 2007]: Improvement from 4.32^k to 4^k. But: "$\Theta(4^k)$"

Balanced subsystems

Drosophila regulatory network



[DasGupta et al., Biosystems 2007]

Balanced graphs

Definition

An undirected graph with edges labeled by $=$ or $\neq$ (signed graph) is balanced iff it contains no cycle with an odd number of $\neq$-edges.

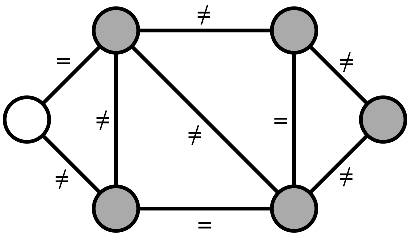
Balanced graphs

Definition

An undirected graph with edges labeled by $=$ or $\neq$ (signed graph) is balanced iff it contains no cycle with an odd number of $\neq$-edges.

Theorem (Kőnig 1936)

A signed graph is balanced iff the vertices can be colored with two colors such that the relation on each edge holds.



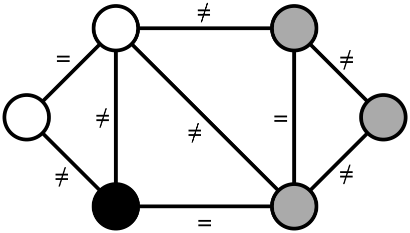
Balanced graphs

Definition

An undirected graph with edges labeled by $=$ or $\neq$ (signed graph) is balanced iff it contains no cycle with an odd number of $\neq$-edges.

Theorem (Kőnig 1936)

A signed graph is balanced iff the vertices can be colored with two colors such that the relation on each edge holds.



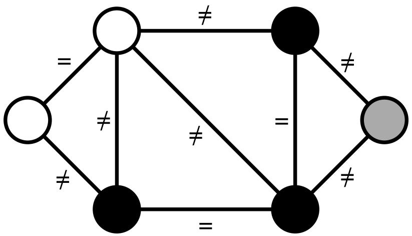
Balanced graphs

Definition

An undirected graph with edges labeled by $=$ or $\neq$ (signed graph) is balanced iff it contains no cycle with an odd number of $\neq$-edges.

Theorem (Kőnig 1936)

A signed graph is balanced iff the vertices can be colored with two colors such that the relation on each edge holds.



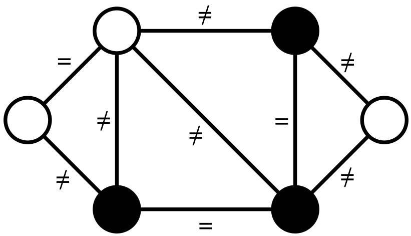
Balanced graphs

Definition

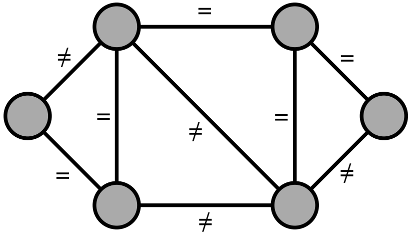
An undirected graph with edges labeled by $=$ or $\neq$ (signed graph) is balanced iff it contains no cycle with an odd number of $\neq$-edges.

Theorem (Kőnig 1936)

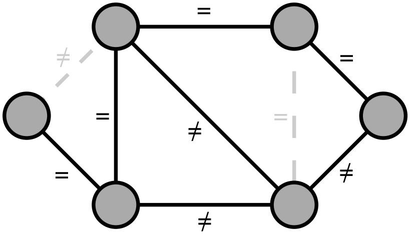
A signed graph is balanced iff the vertices can be colored with two colors such that the relation on each edge holds.



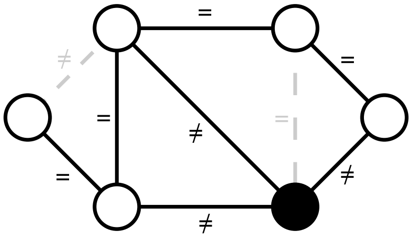
Balanced Subgraph



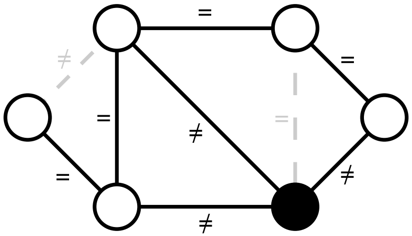
Balanced Subgraph



Balanced Subgraph



Balanced Subgraph



Definition (Balanced Subgraph)

Input: A graph with edges labeled by $=$ or $\neq$.
Task: Find a minimum set of edges to delete such that the graph becomes balanced.

Applications of Balanced Subgraph

- “Monotone subsystems” in gene regulatory networks

  [DasGupta et al., WEA 2006]
- Balance in social networks

  [Harary, Mich. Math. J. 1953]

  e. g. Harary: *A structural analysis of the situation in the Middle East in 1956*, J. Conflict Resolution 1961
- Minimum energy state of magnetic materials (spin glasses)

  [Kasteleyn, J. Math. Phys. 1963]
- Stability of fullerenes

  [Došlić & Vukičević, Discr. Appl. Math. 2007]
- Integrated circuit design

  [Chiang et al., IEEE Trans. CAD of IC & Sys. 2007]

Balanced Subgraph: known results

- BALANCED SUBGRAPH is NP-hard, since it is a generalization of MAX-CUT (MAX-CUT is the special case where all edges are $\neq$)
- A solution that keeps at least 87.8 % of the edges can be found in polynomial time

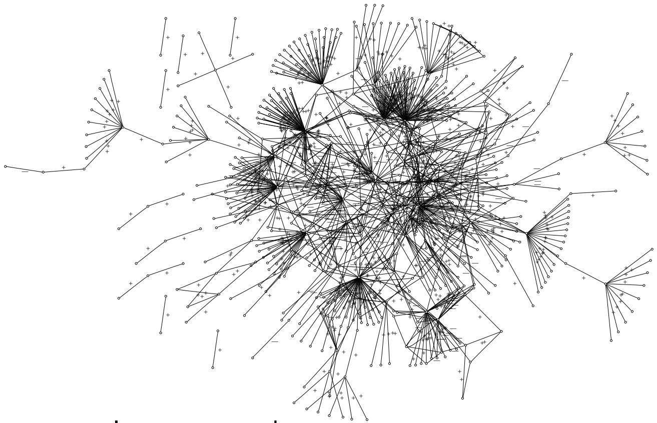
  [THAGARD & VERBEURGT, Cogn. Sci. 1998]
- A solution that deletes at most c times the edges that need to be deleted can probably not be found in polynomial time

  [KHOT, STOC 2002]

Graph structure

Idea

Exploit the structure of the relevant networks



Yeast gene regulatory network

Data reduction

Data reduction

Replace the instance in polynomial time by a simpler, equivalent one.

Data reduction

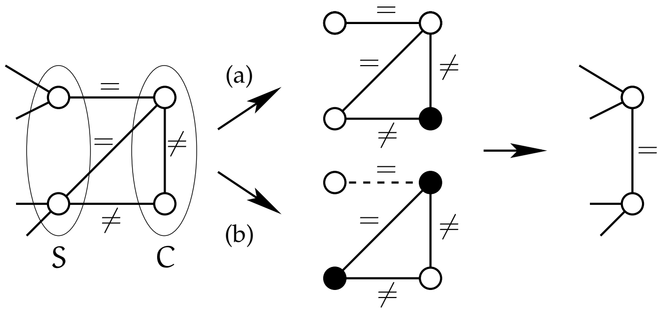
Data reduction

Replace the instance in polynomial time by a simpler, equivalent one.

Example

Delete all degree-1 vertices.

Vertex cut-based data reduction



Data reduction scheme

Data reduction scheme

- Find cut S that cuts off small component C
- For each of the (up to symmetry) $2^{|S|-1}$ colorings of S, determine the size of an optimal solution for $G[S \cup C]$
- Replace in G the subgraph $G[S \cup C]$ by an equivalent smaller gadget

Data reduction scheme

Data reduction scheme

- Find cut S that cuts off small component C
- For each of the (up to symmetry) $2^{|S|-1}$ colorings of S, determine the size of an optimal solution for $G[S \cup C]$
- Replace in G the subgraph $G[S \cup C]$ by an equivalent smaller gadget

Subsumes all 8 data reduction rules given by [Wernicke, 2003] for Edge Bipartization

Filling in the data reduction scheme

- Need to restrict both $|S|$ and $|C|$: we use $|S| \leq 4$ and $|C| \leq 32$

Filling in the data reduction scheme

- Need to restrict both $|S|$ and $|C|$: we use $|S| \leq 4$ and $|C| \leq 32$
- How to construct gadgets that behave equivalently to $S \cup C$?

Gadget construction

Idea

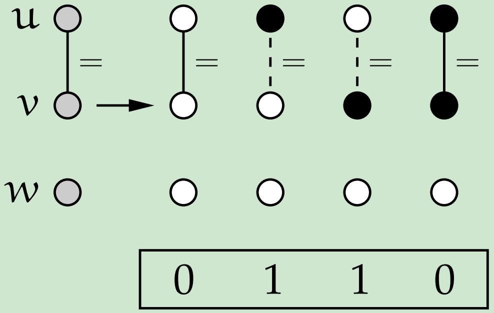
Use atomic gadgets and describe their effect by cost vectors.

Gadget construction

Idea

Use atomic gadgets and describe their effect by cost vectors.

Example



Gadget construction

Theorem

With 10 atomic gadgets, we can emulate the behavior of any component behind a 3-vertex cut.

Gadget construction

Theorem

With 10 atomic gadgets, we can emulate the behavior of any component behind a 3-vertex cut.

Theorem

All cuts with $|S| = 2$ and $|C| \geq 1$ and and all cuts with $|S| = 3$ and $|C| \geq 2$ are subject to data reduction.

Gadget construction

Theorem

With 10 atomic gadgets, we can emulate the behavior of any component behind a 3-vertex cut.

Theorem

All cuts with $|S| = 2$ and $|C| \geq 1$ and and all cuts with $|S| = 3$ and $|C| \geq 2$ are subject to data reduction.

- 4-cuts: 2948 atomic gadgets

Gadget construction

Problem

How to determine an appropriate set of atomic cost vectors for a given cost vector?

Gadget construction

Problem

How to determine an appropriate set of atomic cost vectors for a given cost vector?

Vector Sum Problem

Given a set S of n vectors of length l with nonnegative integer components and a target vector t of length l, find a sub-(multi)-set of vectors from S that sums to t.

Gadget construction

Problem

How to determine an appropriate set of atomic cost vectors for a given cost vector?

Vector Sum Problem

Given a set S of n vectors of length l with nonnegative integer components and a target vector t of length l, find a sub-(multi)-set of vectors from S that sums to t.

- “Equality-constrained multidimensional knapsack”

Gadget construction

Problem

How to determine an appropriate set of atomic cost vectors for a given cost vector?

Vector Sum Problem

Given a set S of n vectors of length l with nonnegative integer components and a target vector t of length l, find a sub-(multi)-set of vectors from S that sums to t.

- "Equality-constrained multidimensional knapsack"
- In our implementation: simple branch & bound

Gadget construction

Problem

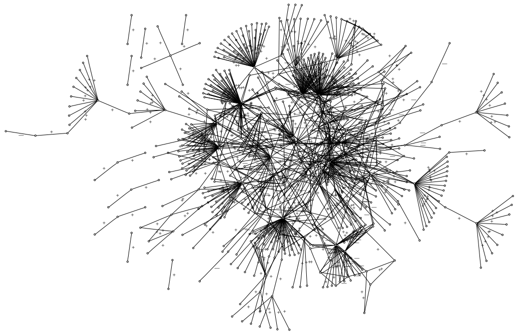
How to determine an appropriate set of atomic cost vectors for a given cost vector?

Vector Sum Problem

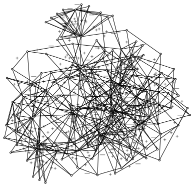
Given a set S of n vectors of length l with nonnegative integer components and a target vector t of length l, find a sub-(multi)-set of vectors from S that sums to t.

- "Equality-constrained multidimensional knapsack"
- In our implementation: simple branch & bound
- Sometimes this is a bottleneck!

Reduction. . . and then?

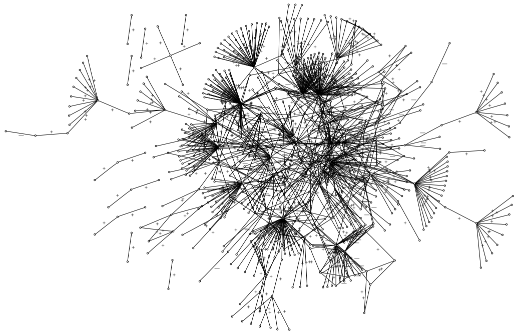


$n = 690,\ m = 1082$

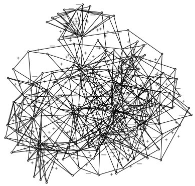


$n = 144,\ m = 405$

Reduction... and then?



$n = 690,\ m = 1082$



$n = 144,\ m = 405$

After data reduction, a hard "core" remains.

Idea

Exploit the fact that only few edges need to be deleted.

Fixed-parameter tractability

Theorem

BALANCED SUBGRAPH *can be solved in* $O(2^k \cdot m)$ *time by a reduction to* EDGE BIPARTIZATION *and using an algorithm based on iterative compression* [GUO et al. 2006].

Fixed-parameter tractability

Theorem

BALANCED SUBGRAPH *can be solved in* $O(2^k \cdot m)$ *time by a reduction to* EDGE BIPARTIZATION *and using an algorithm based on iterative compression* [GUO et al. 2006].

A heuristic speedup trick can give large speedups over this worst-case running time.

Experimental results

			Approximation			Exact alg.	
Data set	n	m	$k \geq$	$k \leq$	t [min]	k	t [min]
EGFR	330	855	196	219	7	210	108
Yeast	690	1082	0	43	77	41	1
Macr.	678	1582	218	383	44	374	1

Experimental results

			Approximation			Exact alg.	
Data set	n	m	$k \geq$	$k \leq$	t [min]	k	t [min]
EGFR	330	855	196	219	7	210	108
Yeast	690	1082	0	43	77	41	1
Macr.	678	1582	218	383	44	374	1

- Yeast is not solvable without reducing 4-cuts
- A real-world network with 688 vertices and 2208 edges could not be solved

Outlook

- Directed case of BALANCED SUBGRAPH (delete minimum number of edges to remove all unbalanced cycles): FPT?
  - Problem: Characterization by two-coloring does not work
- The data reduction scheme is applicable to all graph problems where a coloring or a subset of the vertices is sought. For example:
  - VERTEX COVER
  - DOMINATING SET
  - 3-COLORING
  - FEEDBACK VERTEX SET

  but: need small cuts (e. g., small-world networks)

Protein complexes

Clique Enumeration

Application

- Analysis of biological, social, and other networks
- Finding complexes in protein interaction networks
- Clustering in data mining

Clique Enumeration

Application

- Analysis of biological, social, and other networks
- Finding complexes in protein interaction networks
- Clustering in data mining

Maximal clique enumeration

- Simple model
- NP-hard
- up to $3^{n/3}$ cliques

Clique Enumeration

Application

- Analysis of biological, social, and other networks
- Finding complexes in protein interaction networks
- Clustering in data mining

Maximal clique enumeration

- Simple model
- NP-hard
- up to $3^{n/3}$ cliques

Isolated cliques

- More specific model
- More efficient enumeration algorithms (FPT)

c-Isolation

Definition (Ito, Iwama & Osumi, ESA 2005)

A vertex set S is called c-isolated if on average the vertices in S have less than c neighbors outside of S.

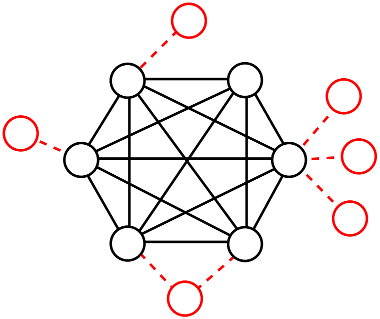
Example: 2-isolation

c-Isolation

Definition (Ito, Iwama & Osumi, ESA 2005)

A vertex set S is called c-isolated if on average the vertices in S have less than c neighbors outside of S.

Example: 2-isolation

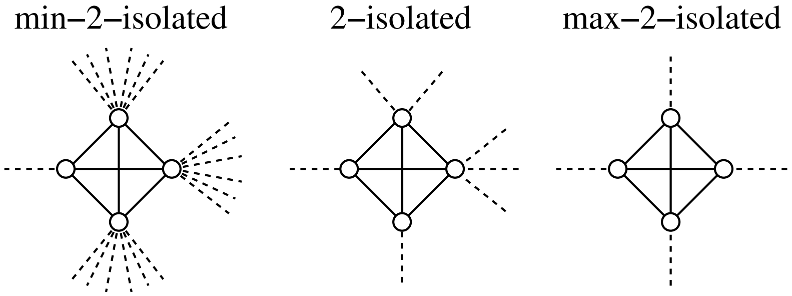


Enumerating maximal c-isolated cliques

Theorem

All maximal c-isolated cliques in a graph G can be enumerated in $O(2.89^c \cdot c^2 m)$ time.

Comparison of isolation concepts



Running times for enumeration of maximal cliques

min-c-isolation	$O(2^c \cdot cm + nm)$
c-isolation	$O(4^c \cdot c^4 m)$
max-c-isolation	$O(2.44^c \cdot cm)$

Finding complexes: Experimental setup

Question

Are isolated cliques a good model for complexes?

Experiment

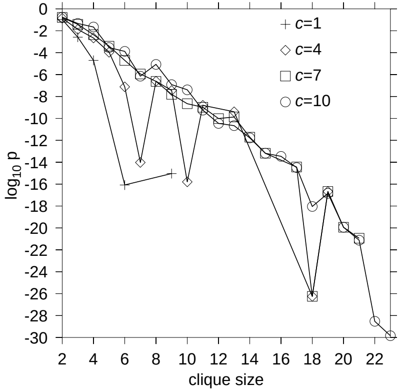
- We retrieved a protein interaction network from BioGRID: *S. cerevisiae*: 5 195 nodes, 70 911 edges.
- We retrieved annotation data for each protein from the Saccharomyces Genome Database (SGD).
- For each enumerated clique, we calculated the *p*-values for the enrichment of annotation terms with the GO Termfinder software, and chose the annotation term with the lowest *p*-value.

Finding complexes: Experimental results

General observations

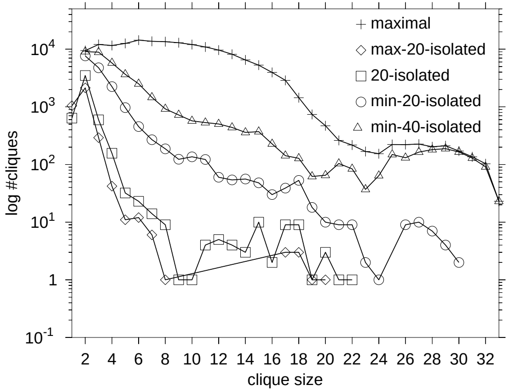
- running time: few seconds
- maximal isolated cliques show more significant enrichment of annotation terms than maximal cliques

Finding complexes: Experimental results



Comparison of mean p-values of the enumerated maximal min-c-isolated cliques and different values of c.

Finding complexes: Experimental results



Distribution of the number of enumerated cliques in the yeast network for different isolation concepts and strengths.

References

- Falk Hüffner, Sebastian Wernicke, and Thomas Zichner:
  Algorithm engineering for color-coding with applications to signaling pathway detection.
  Algorithmica.
  Accepted for publication, June 2007.
- Falk Hüffner, Nadja Betzler, and Rolf Niedermeier:
  Optimal edge deletions for signed graph balancing.
  Proc. 6th Workshop on Experimental Algorithms (WEA '07), 2007.
  LNCS 4525, pp. 297–310.
- Falk Hüffner:
  Algorithms and Experiments for Parameterized Approaches to Hard Graph Problems.
  PhD thesis, Universität Jena, to appear.