An (almost complete) state of the art around the

GRAPH MOTIF problem

Florian Sikora
Université Paris-Est, LIGM - UMR CNRS 8049, France
florian.sikora@univ-mlv.fr

March 30, 2012

Abstract
Here is a (tentative) resume of results around the GRAPH MOTIF problem. For any mistakes, missing results, or comments on this document, please feel free to contact me!

Contents

1 Notations 3

2 The Graph Motif problem 3
  2.1 Network is a tree .................................................. 3
  2.2 Network is a graph .................................................. 4

3 The List-colored Graph Motif problem 4

4 The Biconnected subgraph problem 5

5 The Bridge-connected subgraph problem 5

6 The Max Motif problem 5
  6.1 Network is a tree .................................................. 5
  6.2 Network is a graph .................................................. 5

7 The Min Add problem 5
  7.1 Network is a tree .................................................. 6
  7.2 Network is a graph .................................................. 6

8 The Min Substitution problem 6
  8.1 Network is a tree .................................................. 6
  8.2 Network is a graph .................................................. 6

9 The Min-CC problem 6
  9.1 Network is a path .................................................. 6
  9.2 Network is a tree .................................................. 7
  9.3 Network is a graph .................................................. 7
1 Notations

- Let $G = (V, E)$ be the vertex colored target network, $n = |V|$, $m = |E|$
- Let $M$ be the motif. Let $k$ be the size of the solution. Let $c$ be number of colors in the motif.
- If $|M| = c$, then $M$ is colorful. (Otherwise, it is a multiset)

2 The Graph Motif problem

**Input**: A vertex-colored graph $G$, a multiset of colors $M$.

**Question**: Does $G$ have a connected subset of vertices whose multiset of colors equals $M$?

2.1 Network is a tree

- NP-Complete [LFS06]
- NP-Complete, even for colorful motifs and for trees of maximum degree 3. [FFHV07]
• Polynomial when the motif is colorful on caterpillars [ABH+10]
• Polynomial when the motif is colorful and each color occurs at most 2 times in G [DFV11 Sik11]
• NP-Complete for colorful motifs and rooted trees of height two [ABH+10]
• NP-Complete for colorful motifs on trees, even if a specific node (a root) is asked [ABH+10]
• W[1]-hard when the parameter is c. [FFHV07]
• No polynomial kernel on comb-graphs [ABH+10]

2.2 Network is a graph
• NP-Complete for motifs with 2 colors, even if G is bipartite with maximum degree 4 [FFHV07]
• NP-Complete for colorful motifs on graphs of diameter two [ABH+10]
• Polynomial-time solvable when c is constant and G has a constant treewidth \(O(n^{2cw+2})\), where w is the treewidth of G [FFHV07]
• For colorful motifs : FPT : \(O(3^k.m)\) [BFKN08]
• For colorful motifs : FPT : \(\tilde{O}(2^k k^2 m)\) time and \(\tilde{O}(kn)\) space [GS10]
• For colorful motifs : FPT : \(O(3^k m N_{ins})\) [BHK+09]
• For colorful motifs but allows multiset in the solution : FPT : \(O(3^k m N_{ins})\) [BHK+09]
• For multiset motifs : FPT : \(O(8^k k n^2)\) [FFHV07]
• For multiset motifs : FPT : \(O(4.32^k k^2 m)\) time, \(O(2.47^k n)\) space [BFKN08]
• For multiset motifs : FPT : \(\tilde{O}(4^k k^2 m)\) time and \(\tilde{O}(kn)\) space [GS10]
• For multiset motifs with deletions and r insertions: FPT : \(\tilde{O}(4^k (k + r)^2 m)\) time and \(\tilde{O}((k + r)n)\) space [GS10]

3 The List-colored Graph Motif problem
A set of colors is associated to each network node

Network is a graph
• For colorful motifs but allows multiset in the solution : FPT : \(O(k!.3^k m N_{ins})\) [BHK+09]
• For multiset motifs : FPT : \(O(10.88^k m)\) [BFKN08]
• For multiset motifs: FPT : \(\tilde{O}(4^k k^2 m)\) time and \(\tilde{O}(kn)\) space (implicit algorithm) [GS10]
4 The Biconnected subgraph problem
The solution must be biconnected instead of connected
- $W[1]$-complete with respect to $k$ [BFKN08]

5 The Bridge-connected subgraph problem
The solution must be bridge connected instead of connected
- $W[1]$-complete with respect to $k$ [BFKN08]

6 The Max Motif problem
Want an maximum sized connected occurrence of $M$ is $G$.
Hardness results of the Graph Motif problem hold since it is a special case of the Max Motif problem. For the same reason, FPT algorithm is unlikely if the parameter is the number of deletions.

6.1 Network is a tree
- APX-Hard even when $G$ is a tree of maximum degree 3, colorful motif and each color occurs at most twice in $G$ [DFV09]
- Not approximable within factor $|V|^{1/2-\epsilon}$, for any $\epsilon > 0$, even if the motif is colorful and each color occurs at most twice in $G$ [RS12]
- Not approximable within factor $2^{\log^2 n}$, for any $\delta < 1$ (equivalent to no constant approximation ratio) even if the motif is colorful [DFV09]
- Colorful Motifs : Exponential algorithm : $O(1.33^n \cdot poly(n))$ [DFV09]
- Multiset Motifs : Exponential algorithm : $O(1.62^n \cdot poly(n))$ [DFV09]
- Multiset motifs : FPT : $O(k2^k n^3 \log n)2^{O(k)}$ [DFV09]

6.2 Network is a graph
- Multiset motifs : FPT : $O(2^{5k}kn^2 \log^2 n)4^{O(k)}$ [DFV09]
- For multiset motifs: FPT : $\tilde{O}(4^k k^2 m)$ time and $\tilde{O}(kn)$ space [GS10]

7 The Min Add problem
Want an occurrence of $M$ in $G$ with the minimum number of insertions. Equivalent to the Graph Motif problem with a bounded number of insertions.
Hardness results of the Graph Motif problem hold since it is a special case of the Min Add problem. For the same reason, FPT algorithm is unlikely if the parameter is the number of additions.
7.1 Network is a tree
- NP-hard, even with $G$ is a tree of max degree 4, the motif is colorful and each color occurs twice in $G$ [DFV11]

7.2 Network is a graph
- For colorful motifs: FPT : $\mathcal{O}(3^k.m.N_{ins})$ [BHK+09]
- For colorful motifs but allows multiset in the solution: FPT : $\mathcal{O}(3^k.m.N_{ins})$ [BHK+09]
- For multiset motifs with deletions and $r$ insertions: FPT : $\tilde{\mathcal{O}}(4^k(k+r)^2m)$ time and $\tilde{\mathcal{O}}((k+r)n)$ space [GS10]

8 The Min Substitution problem
Want an occurrence of $M$ in $G$ with the minimum number of substitutions

Hardness results of the Graph Motif problem hold since it is a special case of the MIN SUBSTITUTION problem. For the same reason, FPT algorithm is unlikely if the parameter is the number of substitutions.

8.1 Network is a tree
- NP-hard, even with $G$ is a tree of max degree 4, the motif is colorful and each color occurs twice in $G$ [DFV11]
- Not approximable within factor $c \log |V|$, $c$ a constant, even if the motif is colorful and $G$ a tree of depth 2. [RST12]
- W[2]-hard when parametrized by the number of substitutions even is $M$ is colorful [RST12]

8.2 Network is a graph
- FPT $\mathcal{O}^*((3e)^{O(k)})$ [DFV11]

9 The Min-CC problem
Want an occurrence of the motif with the minimum number of connected components in the solution

Hardness results of the Graph Motif problem hold since it is a special case of the MIN-CC problem. For the same reason, FPT algorithm is unlikely if the parameter is the number of CC.

9.1 Network is a path
- Polynomial-time solvable if $c$ is a constant in $\mathcal{O}(n^{c+4})$ [DFV07]
- W[2]-hard when parametrized by the number of connected components [BFKN08]
• APX-Hard even for colorful motifs, each color appears exactly twice in $G$

9.2 Network is a tree
• Solvable in $O(n^2 2^{\frac{4k}{c}})$ [DFV07]
• FPT in $O(n^2 k^{(c+1)^2+1})$ [DFV07]
• W[1]-hard when parametrized by $c$ [FFHV07]
• W[2]-hard when parametrized by the number of connected components even is $M$ is colorful [DFV07]
• Not approximable within $c \log n$ for a constant $c > 0$ even if $M$ is colorful [DFV07]

9.3 Network is a graph
• For colorful motifs (search for $r$ connected components): FPT : $\tilde{O}(2^k k^2 r^2 m)$
  time and $\tilde{O}(k r n)$ space [GS10]
• FPT by $k$ [DFV07]
• FPT in $O(|\ln(\epsilon)| \cdot 4.32^k k^2 m)$ [BFKN08]
• For multiset motifs (search for $r$ connected components): FPT : $\tilde{O}(4^k k^2 r^2 m)$
  time and $\tilde{O}(k r n)$ space [GS10]

10 The Edge-Weighted Graph Motif problem
$G$ is weighted on the edges.

10.1 Network is a graph
• These three results want to minimize the weight of the edge-cut between
  the solution and the rest of the graph.
• Multiset motifs : FPT : $O(|\log(\epsilon)| 2^{k+k d} n)$, with $d$ maximum degree of $G$
  [BRS09]
• Multiset motifs : Branch and bound algorithm : $O(m \log(m) + n^b)$, with
  $b$ maximum number of bounds [BRS09]
• Multiset motifs : FPT : $O(m. c^k. 2^w. (m. k. 2^k. 3^w + \omega + d))$ time, $O(m. c^k. 2^w)$
  space, with $d$ maximum degree of $G$, $w$ treewidth of $G$ [BRS09]
• These two results want to minimize the weight of the edges in the solution.
• For colorful motifs (sum of weight in the solution < $r$): FPT : $\tilde{O}(2^k k^2 r^2 m)$
  time and $\tilde{O}(k r n)$ space [GS10]
• For multiset motifs (sum of weight in the solution < $r$): FPT : $\tilde{O}(4^k k^2 r^2 m)$
  time and $\tilde{O}(k r n)$ space [GS10]
11 The #Graph Motif problem

Network is a graph

- For colorful motifs: FPT: $O(2^k k^3 m)$ time and $O(k^2 n)$ space [GS10]
- For multiset motifs: #W[1]-hard for parameter $k$, even with two colors in the motif [GS10]

12 The Constrained Graph Motif problem

Given a set of mandatory vertices $V_M \subseteq V$, find an occurrence of $M$ s.t. all the mandatory vertices are in the solution.

12.1 Network is a graph

- FPT with the same complexity of the Graph Motif problem, when the size of the solution is the parameter [DFV11]
- FPT with parameter $t = |M| - |V_M|$ if $G$ is of bounded treewidth [DFV11] $O(w^3 w n^{2.44271})$, where $w$ is the treewidth of the graph [DFV11]
- W[2]-hard if the parameter is $t = |M| - |V_M|$, even when $G$ is a graph of diameter 2 [DFV11]

13 The Module Graph Motif problem

The solution is $G$ must be a graph module (instead of being simply connected)

13.1 Network is a forest

- NP-Complete even if the motif is colorful and $G$ is a collection of paths of length 3 [RS12]

13.2 Network is a graph

- Polynomial if the module must be strong [RS12]
- FPT in $O(2^k |V|^2)$ time and $O(2^k |V|)$ space, where $k$ is the size of the solution [RS12]
- FPT if parameterized by $(k, |C|)$, where $C$ is the set of colors [RS12]
- FPT if a list of colors is given for each node [RS12]

14 Softwares

- Motus [LPS06] [http://genome.crg.es/~vlacroix/motus/]
• Torque \cite{BHK10}. \url{http://www.cs.tau.ac.il/~bnet/torque.html} Internet server, dynamic programming + Integer Linear Programming (CPLEX). Only colorful motifs. Allow insertions and deletions. Allow list colored graph motif. Only one solution.

• GraMoFoNe \cite{BSV10}. \url{http://igm.univ-mlv.fr/AlgoB/gramofone/} Cytoscape java plugin. Pseudo Boolean Programming. Multiset motifs, allow insertions and deletions. Allow list colored graph motif.

References


