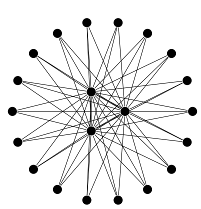
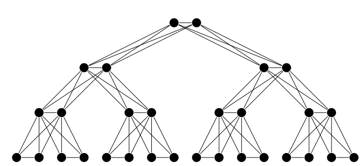
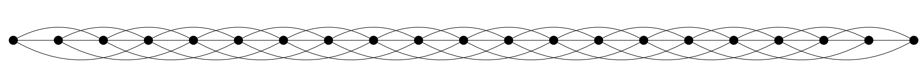
Technika „tnij i zliczaj” w grafowych problemach spójności parametryzowanych szerokością drzewiastą

Autoreferat rozprawy doktorskiej

Marek Cygan

Wprowadzenie W powszechnym przekonaniu dla żadnego problemu NP-trudnego nie istnieje algorytm wielomianowy. Jednakże wiele grafowych problemów NP-trudnych możemy rozwiązać w czasie wielomianowym dla wybranych klas grafów takich jak drzewa, grafy zewnętrznie planarne czy też grafy szeregowo-równoległe (ang. series-parallel).

Szerokość drzewiasta (ang. treewidth) jest parametrem grafowym wprowadzonym niezależnie przez Rosego w roku 1974 [23] (pod nazwą częściowych k-drzew) oraz przez Robertsona i Seymoura [22] w roku 1984. Parametr ten ma za zadanie określać stopień trudności danej instancji, przy rozwiązywaniu problemów grafowych, mierząc stopień podobieństwa grafu do drzewa. Definicja grafów o ograniczonej szerokości drzewiastej jest zbyt złożona aby ją w tym miejscu przytoczyć. Możemy jednak nadmienić że drzewa mają szerokość drzewiastą równą jeden, grafy zewnętrznie planarne oraz szeregowo-równoległe nie większą niż dwa, a jako przykład prezentujemy trzy grafy o szerokości drzewiastej nie większej niż 4, które ilustrują strukturę grafów o ograniczonej szerokości drzewiastej.

Jedną z cech grafów o ograniczonej szerokości drzewiastej jest zbiór separatorów, który umożliwia efektywne rozwiązywanie trudnych problemów obliczeniowych za pomocą programowania dynamicznego. W szczególności Courcelle [6] pokazał algorytm, który mając daną formułę ϕ w monadycznej logice drugiego rzędu oraz graf $G = (V, E)$ o szerokości drzewiastej nie większej niż t sprawdza czy formuła ϕ jest prawdziwa w grafie G w czasie $O(f(|\phi|, t)(|V| + |E|))$. Zauważmy że dla ustalonej formuły oraz $t = O(1)$ jest to algorytm liniowy. Za pomocą odpowiednich formuł możemy modelować większość naturalnych problemów grafowych. Niestety funkcja f może być wielokrotnie wykładnicza i z tego powodu opracowuje się algorytmy dla poszczególnych problemów grafowych z osobna.

Niektóre problemy grafowe mają charakter lokalny ze względu na swoją definicję. Przykładowo w problemie pokrycia wierzchołkowego możemy sprawdzić czy dany zbiór wierzchołków jest rozwiązaniem na podstawie weryfikacji każdej krawędzi grafu z osobna. Dla wielu problemów lokalnych takich jak problem pokrycia wierzchołkowego, zbioru dominującego, czy też zliczania liczby doskonałych skojarzeń nietrudno uzyskać algorytmy o złożoności $c^{\mathrm{tw}(G)}|V|^{O(1)}$, dla pewnej stałej c, gdzie $\mathrm{tw}(G)$ jest szerokością drzewiastą grafu G.

Jednakże dla znaczącej grupy problemów definicja rozwiązania posiada wymóg dotyczący całej struktury grafu, tak jak w problemie cyklu Hamiltona, w którym chcemy aby cykl przechodził przez każdy wierzchołek grafu dokładnie jeden raz lub też w problemie spójnego pokrycia wierzchołkowego, w którym zbiór wierzchołków aby był rozwiązaniem nie tylko powiniem być pokryciem wierzchołkowym, ale również powiniem indukować podgraf spójny. Dla problemów z globalnym wymogiem dotychczasowe najszybsze algorytmy mają złożoność $c^{\mathrm{tw}(G)\log \mathrm{tw}(G)}|V|^{O(1)}$, co wynika z liczby różnych częściowych rozwiązań przechodzących przez separator (lub innymi słowy z liczby stanów w programowaniu dynamicznym). Marx, Lokshtanov i Saurabh [17] opracowali schemat dowodzenia optymalności istniejących algorytmów dla problemów lokalnych w grafach o ograniczonej szerokości drzewiastej (przy silnych założeniach teoriozłożonościowych). Jednocześnie w pracy [17] postawiono pytanie czy da się wykazać, że istniejące algorytmy o złożoności $c^{\mathrm{tw}(G)\log \mathrm{tw}(G)}|V|^{O(1)}$, dla problemów z globalnym wymogiem spójności takich jak problem spójnego pokrycia wierzchołkowego czy też problem cyklu Hamiltona, są optymalne.

Główne wyniki rozprawy Przedmiotem rozprawy jest nowa technika, nazwana *tnij i zliczaj*, mająca zastosowanie w wielu problemach z globalnym wymogiem spójności dla grafów o ograniczonej szerokości drzewiastej. W technice tej sprowadzamy dany problem globalny do nowego problemu lokalnego, w którym należy wyznaczyć liczbę „przeciętych" obiektów modulo dwa poprzez:

- zagwarantowanie, że niespójne obiekty będą występować parzystą liczbę razy, natomiast spójne obiekty będą policzone nieparzystą liczbę razy,
- zastosowanie lematu o izolacji w celu zagwarantowania, że jeśli oryginalny problem decyzyjny miał rozwiązanie, to liczba jego rozwiązań jest nieparzysta.

Przykładowo problem cyklu Hamiltona sprowadzamy do problemu zliczania pokryć cyklowych, w których każdy wierzchołek wybiera stronę lewą lub stronę prawą i każda z krawędzi pokrycia cyklowego łączy wierzchołki będące po tej samej stronie. Istota techniki polega na tym, że zdefiniowane „przecięte" obiekty można zliczać za pomocą klasycznych metod programowania dynamicznego, stosowanych w problemach lokalnych, w czasie $c^{\mathrm{tw}(G)}|V|^{O(1)}$.

Stosując technikę tnij i zliczaj otrzymałem następujące algorytmy:

- dla problemu cyklu Hamiltona, o złożoności $4^{\mathrm{tw}(G)}|V|^{O(1)}$,
- dla problemu spójnego pokrycia wierzchołkowego, o złożoności $3^{\mathrm{tw}(G)}|V|^{O(1)}$,
- dla problemu zbioru rozcyklającego (ang. feedback vertex set), o złożoności $3^{\mathrm{tw}(G)}|V|^{O(1)}$,
- dla problemu spójnego zbioru dominującego, o złożoności $4^{\mathrm{tw}(G)}|V|^{O(1)}$,
- dla problemu drzewa Steinera, o złożoności $3^{\mathrm{tw}(G)}|V|^{O(1)}$.

Ponadto używając technik z pracy [17] pokazałem, że przy silnych założeniach teoriozłożonościowych, dla wszystkich wyżej wymienionych problemów poza cyklem Hamiltona uzyskana złożoność jest optymalna. Innymi słowy technika tnij i zliczaj nie tylko pozwala na usunięcie czynnika logarytmicznego z wykładnika ale również prowadzi do uzyskania optymalnej podstawy funkcji wykładniczej.

Rozszerzając technikę tnij i zliczaj o użycie zdefiniowanych przeze mnie *znaczników* pokazałem, że w czasie $c^{\mathrm{tw}(G)}|V|^{O(1)}$ nie tylko możemy kontrolować spójność konstruowanego rozwiązania ale również potrafimy rozwiązywać problemy, w których rozwiązanie może się składać z wielu spójnych składowych, przy założeniu że rozwiązanie powinno minimalizować liczbę spójnych składowych. Dzięki temu uzyskałem algorytm o złożoności $6^{\mathrm{tw}(G)}|V|^{O(1)}$ dla problemu pokrycia grafu skierowanego za pomocą minimalnej liczby cykli.

Naturalnym kierunkiem badań było sprawdzenie czy dla problemów, w których należy zmaksymalizować liczbę spójnych składowych rozwiązania (tak jak w problemie pakowania cykli) również da się otrzymać algorytmy o złożoności $c^{\mathrm{tw}(G)}|V|^{O(1)}$. W rozprawie doktorskiej wykazałem jednak, że dla problemu pakowania cykli (zarówno w wariancie skierowanym jak i nieskierowanym) nie da się otrzymać algorytmów o złożoności $c^{o(\mathrm{tw}(G)\log \mathrm{tw}(G))}|V|^{O(1)}$ przy założeniu, że nie istnieje podwykładniczy algorytm dla problemu spełnialności formuł. Wynik ten wskazuje istotną różnicę pomiędzy globalnymi problemami, w których liczba spójnych składowych rozwiązania powinna być jak najmniejsza a tymi, w których liczba spójnych składowych rozwiązania powinna być jak największa.

Konsekwencje Jako że algorytmy dla grafów o ograniczonej szerokości drzewiastej często pojawiają się jako podprocedury w algorytmach aproksymacyjnych, algorytmach parametryzowanych oraz dokładnych algorytmach wykładniczych, technika tnij i zliczaj pozwala na uzyskanie lepszych złożoności w tych dziedzinach. Opiszemy teraz dwa moim zdaniem najciekawsze zastosowania prezentowanej w rozprawie techniki.

Łącząc technikę tnij i zliczaj z metodą iterowanej kompresji [21] dla problemów parametryzowanych wielkością rozwiązania uzyskałem następujące ulepszenia.

- Problem zbioru rozcyklającego (ang. feedback vertex set) ma długą historię algorytmów parametryzowanych [1, 3, 5, 10, 11, 12, 14, 16, 19, 20]. Uprzednio najlepszy algorytm pochodził od Cao, Chena i Liu i miał złożoność $3.83^k|V|^{O(1)}$ [4]. W rozprawie prezentuję algorytm o złożoności $3^k|V|^{O(1)}$.
- Dla problemu spójnego pokrycia wierzchołkowego poprawiam algorytm Binkele-Raiblego [2] o złożoności $2.4882^k|V|^{O(1)}$ do $2^k|V|^{O(1)}$.
- Dla problemu spójnego zbioru rozcyklającego poprawiam algorytm Misry i innych [18] o złożoności $46.2^k|V|^{O(1)}$ do $3^k|V|^{O(1)}$.

Korzystając z twierdzenia Fomina i innych [13] które dowodzi, że grafy o ograniczonym stopniu mają ograniczoną szerokość drzewiastą, w połączeniu z techniką tnij i zliczaj otrzymałem najszybszy algorytm dla problemu cyklu Hamiltona w grafach kubicznych poprawiając algorytm Iwamy i Nakashimy [15] o złożoności $O(1.252^n)$ do złożoności $O(1.201^n)$.

Redukcje Jak wspomniano powyżej w rozprawie zaprezentowałem najszybszy obecnie znany algorytm paramametryzowany wielkością rozwiązania dla problemu spójnego pokrycia wierzchołkowego. Algorytm ten nie tylko rozwiązuje problem decyzyjny ale również potrafi znaleźć parzystość liczby rozwiązań w czasie $2^k|V|^{O(1)}$. Okazuje się, że przy silnych założeniach teorio-złożonościowych nie istnieje $\varepsilon > 0$ oraz algorytm o złożoności $(2-\varepsilon)^k|V|^{O(1)}$, który oblicza parzystość liczby spójnych pokryć wierzchołkowych wielkości k. Zgodnie z naszą najlepszą wiedzą zaprezentowany w pracy algorytm jest pierwszym dla problemu parametryzowanego wielkością rozwiązania dla którego istnieją silne przesłanki optymalności podstawy funkcji wykładniczej.

Uwagi końcowe Wyniki zawarte w rozprawie zostały uzyskane przy następującej współpracy.

- Technika tnij i zliczaj jak również zdecydowana większość zawartości merytorycznej mojej rozprawy pochodzi z pracy [8] opracowanej razem z Jesperem Nederlofem, Marcinem Pilipczukiem, Michałem Pilipczukiem, Johanem van Rooijem oraz Jakubem Onufrym Wojtaszczykiem, opublikowanej na konferencji FOCS 2011.

- Metoda efektywnego obliczania uogólnionego splotu podzbiorów używana jako narzędzie przy niektórych zastosowaniach techniki tnij i zliczaj pochodzi z pracy [9] opracowanej razem z Marcinem Pilipczukiem, opublikowanej na konferencji ICALP 2009.

- Optymalność algorytmu parametryzowanego wielkością rozwiązania dla problemu spójnego pokrycia wierzchołkowego pochodzi z pracy [7] opracowanej razem z Holgerem Dellem, Danielem Lokshtanovem, Dánielem Marxem, Jesperem Nederlofem, Yoshio Okamoto, Ramamohanem Paturim, Saketem Saurabhem oraz Magnusem Wahlströmem. Praca ta jest aktualnie w recenzji.

Literatura

[1] A. Becker, R. Bar-Yehuda, D. Geiger. Randomized algorithms for the loop cutset problem. *Journal of Artificial Intelligence Research*, 12(1):219–234, 2000.

[2] D. Binkele-Raible. *Amortized Analysis of Exponential Time and Parameterized Algorithms: Measure and Conquer and Reference Search Trees.* Praca doktorska, University of Trier, Trier, Germany, 2010.

[3] H. L. Bodlaender. On disjoint cycles. *International Journal of Foundations of Computer Science*, 5(1):59–68, 1994.

[4] Y. Cao, J. Chen, Y. Liu. On feedback vertex set new measure and new structures. H. Kaplan, redaktor, *12th Scandinavian Symposium and Workshops on Algorithm Theory, SWAT 2010*, wolumen 6139 serii *Lecture Notes in Computer Science*, strony 93–104. Springer, 2010.

[5] J. Chen, F. V. Fomin, Y. Liu, S. Lu, Y. Villanger. Improved algorithms for feedback vertex set problems. *Journal of Computer and System Sciences*, 74(7):1188–1198, 2008.

[6] B. Courcelle. The Monadic Second-Order Logic of Graphs. I. Recognizable Sets of Finite Graphs. *Inf. Comput.*, 85(1):12–75, 1990.

[7] M. Cygan, H. Dell, D. Lokshtanov, D. Marx, J. Nederlof, Y. Okamoto, R. Paturi, S. Saurabh, M. Wahlström. On problems as hard as CNF-SAT. *Manuskrypt*, 2011.

[8] M. Cygan, J. Nederlof, M. Pilipczuk, M. Pilipczuk, J. M. M. van Rooij, J. O. Wojtaszczyk. Solving connectivity problems parameterized by treewidth in single exponential time. *52th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2011*. IEEE Computer Society, 2011.

[9] M. Cygan, M. Pilipczuk. Exact and approximate bandwidth. S. Albers, A. Marchetti-Spaccamela, Y. Matias, S. E. Nikoletseas, W. Thomas, redaktorzy, *36th International Colloquium on Automata, Languages and Programming (1), ICALP 2009*, wolumen 5555 serii *Lecture Notes in Computer Science*, strony 304–315. Springer, 2009.

[10] F. K. H. A. Dehne, M. R. Fellows, M. A. Langston, F. A. Rosamond, K. Stevens. An $O(2^{O(k)}n^3)$ FPT algorithm for the undirected feedback vertex set problem. L. Wang, redaktor, *13th Annual International Conference on Computing and Combinatorics, COCOON 2005*, wolumen 3595 serii *Lecture Notes in Computer Science*, strony 859–869. Springer, 2005.

[11] R. G. Downey, M. R. Fellows. Fixed parameter tractability and completeness. K. Ambos-Spies, S. Homer, U. Schöning, redaktorzy, *Complexity Theory: Current Research*, strony 191–225. Cambridge University Press, 1993.

[12] R. G. Downey, M. R. Fellows. *Parameterized Complexity*. Springer, 1999.

[13] F. V. Fomin, S. Gaspers, S. Saurabh, A. A. Stepanov. On two techniques of combining branching and treewidth. *Algorithmica*, 54(2):181–207, 2009.

[14] J. Guo, J. Gramm, F. Hüffner, R. Niedermeier, S. Wernicke. Compression-based fixed-parameter algorithms for feedback vertex set and edge bipartization. *Journal of Computer and System Sciences*, 72(8):1386–1396, 2006.

[15] K. Iwama, T. Nakashima. An improved exact algorithm for cubic graph TSP. G. Lin, redaktor, *13th Annual International Conference on Computing and Combinatorics, COCOON 2007*, wolumen 4598 serii *Lecture Notes in Computer Science*, strony 108–117. Springer, 2007.

[16] I. A. Kanj, M. J. Pelsmajer, M. Schaefer. Parameterized algorithms for feedback vertex set. R. G. Downey, M. R. Fellows, F. K. H. A. Dehne, redaktorzy, *1st International Workshop on Parameterized and Exact Computation, IWPEC 2004*, wolumen 3162 serii *Lecture Notes in Computer Science*, strony 235–247. Springer, 2004.

[17] D. Lokshtanov, D. Marx, S. Saurabh. Known algorithms on graphs of bounded treewidth are probably optimal. *22st Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011*, 2011. Accepted for publication, to appear.

[18] N. Misra, G. Philip, V. Raman, S. Saurabh, S. Sikdar. FPT algorithms for connected feedback vertex set. M. S. Rahman, S. Fujita, redaktorzy, *4th International Workshop on Algorithms and Computation, WALCOM 2010*, wolumen 5942 serii *Lecture Notes in Computer Science*, strony 269–280. Springer, 2010.

[19] V. Raman, S. Saurabh, C. R. Subramanian. Faster fixed parameter tractable algorithms for undirected feedback vertex set. P. Bose, P. Morin, redaktorzy, *13th International Symposium on Algorithms and Computation, ISAAC 2002*, wolumen 2518 serii *Lecture Notes in Computer Science*, strony 241–248. Springer, 2002.

[20] V. Raman, S. Saurabh, C. R. Subramanian. Faster fixed parameter tractable algorithms for finding feedback vertex sets. *ACM Transactions on Algorithms*, 2(3):403–415, 2006.

[21] B. A. Reed, K. Smith, A. Vetta. Finding odd cycle transversals. *Operations Research Letters*, 32(4):299–301, 2004.

[22] N. Robertson, P. D. Seymour. Graph minors. III. Planar tree-width. *Journal of Combinatorial Theory, Series B*, 36(1):49–64, 1984.

[23] D. J. Rose. On simple characterizations of k-trees. *Discrete Mathematics*, 7(3-4):317–322, 1974.