

<https://doi.org/10.7236/JIIBC.2022.22.5.1>
JIIBC 2022-5-1

최대 매칭 문제의 최소차수 정점 간 간선 선택 알고리즘

Algorithm for Minimum Degree Inter-vertex Edge Selection of Maximum Matching Problem

이상운*

Sang-Un, Lee *

요약 본 논문은 최대 매칭 문제(MCM)를 다루었다. MCM은 일반적으로 증대경로 기법으로 구한다. 일반 그래프에 대한 MCM을 구하는 증대경로 알고리즘으로는 $O(\sqrt{nm})$ 복잡도, 이분 그래프에 대해서는 $O(m \log n)$ 복잡도를 갖고 있다. 반면에, 본 논문에서는 주어진 그래프가 일반 그래프나 이분그래프의 그래프 종류에 상관없이 항상 $O(n)$ 복잡도로 MCM을 구하는 알고리즘을 제안하였다. 제안된 알고리즘은 “최대 매칭을 구하기 위해서는 가능한 많은 정점 쌍의 간선을 선택해야만 한다.”는 기본 원리에 근거하여 최소차수 정점 u 와 $N_G(u)$ 들 중 최소차수 정점 v 간 간선 $\{u, v\}$ 를 $\nu(G) = k$ 회 단순히 선택하는 간단한 방법이다. 제안된 알고리즘을 일반그래프와 이분그래프의 다양한 실험 데이터들에 적용한 결과 $\nu(G)$ 를 정확하게 구할 수 있음을 보였다.

Abstract This paper deals with the maximum cardinality matching(MCM) problem. The augmenting path technique is well known in MCM. MCM is obtained by $O(\sqrt{nm})$ time complexity augmenting path algorithm for the general graph, and $O(m \log n)$ algorithm for the bipartite graph. On the other hand, this paper suggests $O(n)$ linear time algorithm. The proposed algorithm based on the basic principle of as possible as largest selected inter-vertex edges in order to obtain the MCM. This paper simply selects edge $\{u, v\}$ that the minimum degree vertex u and minimum degree vertex v in $N_G(u)$ $\nu(G) = k$ times iteration. For various general and bipartite graphs experimental data, this algorithm can be get the $\nu(G)$ exactly.

Key Words : Maximum cardinality matching, Maximal, Maximum, minimum degree, Neighborhood

1. 서론

일반 그래프 $G=(V,E), n=|V|, m=|E|$ 와 이분 그래프 $K_{m,n} (m \leq n)$ 에서 최대 매칭(maximum matching 또는 maximum cardinality matching, MCM)을 구하

는 문제는 이분그래프를 활용하는 빅 데이터로부터 유용한 정보를 추출하는 데이터마이닝이나 작업할당 문제에 서 중요한 부분을 차지하고 있다.^[1]

매칭이란 두 정점 u 와 v 가 인접하여 간선 $\{u, v\}$ 가 존재할 경우 간선 $\{u, v\}$ 를 1:1로 선택함으로써 정점 u 와 v

*정회원, 강릉원주대학교 과학기술대학 멀티미디어공학과
접수일자 : 2022년 3월 9일, 수정완료 : 2022년 9월 9일
게재확정일자 : 2022년 10월 7일

Received: 9 March, 2022 / Revised: 9 September, 2022 /
Accepted: 7 October, 2022

*Corresponding Author: sulee@gwnu.ac.kr

Dept. of Multimedia Eng., Gangneung-Wonju National University, Korea

를 매칭시킴을 의미한다. 따라서 일단 $\{u, v\}$ 간선을 매칭시키면 u 와 v 에 부속된 나머지 간선들은 삭제되어야만 한다.^[2]

이와 같이 매칭 M 의 간선 수 $\nu(G) = k$ 를 최대로 할 수 있는 경우를 MCM이라 하며, 일반 그래프는 최대 $\lfloor n/2 \rfloor, K_{m,n} (m \leq n)$ 이분그래프는 최대 m 이다.

MCM을 구하는 알고리즘으로는 정확한 해를 다항시간으로 구할 수 있는 알고리즘이 알려져 있지 않다는 가정 하에 컴퓨터 프로그램을 활용해 다항시간으로 근사해를 구하는 메타휴리스틱 기법과 다항시간으로 정확한 해를 구하는 규칙이 존재한다는 휴리스틱 알고리즘으로 분류될 수 있다.

메타휴리스틱 기법의 일종인 진화 알고리즘(evolutionary algorithm, EA)으로는 Giel과 Wegener^[3], He와 Tao^[4]가 있으며, 유전자 알고리즘(genetic algorithm, GA)으로는 Shetter와 Angadi^[5]가 있다.

휴리스틱 알고리즘의 일종인 증대경로법으로는, Hofcroft와 Karp^[6]의 $O(n^{5/2})$ 알고리즘, Micali와 Vazirani^[7]의 $O(\sqrt{nm})$ 알고리즘, Blum^[8]의 $O(\sqrt{nm})$ 알고리즘, Witzgall과 Zahn^[2]의 $O(n^3)$ Edmond's 알고리즘, Kameda와 Munro^[9]의 $O(mn)$ 알고리즘이 있으며, 우선순위 큐를 활용한 Galil^[10]의 $O(m \log n)$ 알고리즘, 분광학습법을 적용한 Witzgall과 Zahn^[11]의 $O(m)$ 알고리즘이 있다.

$G = (V, E), n = |V|, m = |E|, m > n$ 의 MCM을 구하는데 있어 본 논문에서는 $O(n)$ 의 선형시간 복잡도를 갖는 알고리즘을 제안한다. 2장에서는 최대도 가능한 매칭(최대한, maximal matching)과 최대 매칭(maximum matching) 개념을 정의하고, 증대경로 기법으로 MCM을 찾는 과정을 논한다. 3장에서는 최소 차수 정점 간 간선을 단순히 $\nu(G) = k$ 회 선택하는 알고리즘을 제안한다. 4장에서는 제안된 알고리즘을 적용하여 MCM을 찾을 수 있는지 검증하여 본다.

II. 용어 정의와 증대경로법

그림 1은 그래프 G_1, G_2, G_3 의 3개 그래프를 보여주고 있다. 이 그래프들에 대한 최대한 매칭(maximal matching)과 최대 매칭(maximum matching) 결과는 그림 2에 제시되어 있다.

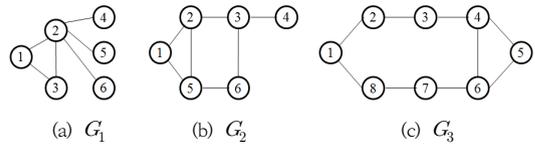


그림 1. 예제 그래프
Fig. 1. Example graphs

[정의 1] 최대한 매칭(maximal matching, MM)이란 매칭 M 에 속하지 않은 어떤 간선이 M 에 추가된다면 더 이상 매칭이 되지 않는 속성을 가진 그래프 G 의 매칭 M 을 의미한다. 다시 말해, 그래프 G 에 있는 모든 간선이 M 에 있는 적어도 하나의 간선과 비공집합 교차(non-empty intersection)를 가진다면 G 의 매칭 M 이 최대한 매칭이라 한다.

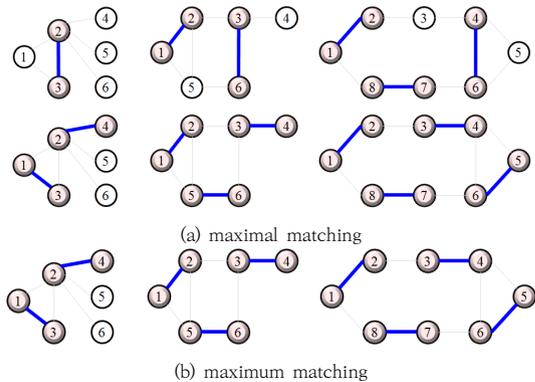


그림 2. 예제 그래프에 대한 최대한과 최대 매칭
Fig. 2. Maximal and maximum matching for example graphs

그래프 G_1 에서 $\{2,3\}$ 간선이 M 에 포함되면 정점 ②와 ③에 부속된 간선으로 연결된 인접하는 정점 ①, ④, ⑤, ⑥은 $\{2,3\}$ 간선이 지배하여 더 이상 M 에 포함시킬 간선이 없게 된다. 따라서 $\{2,3\}$ 은 MM이다. 마찬가지로 $M = \{\{1,3\}, \{2,4\}\}$ 인 경우도 MM이 된다.

[정의 2] 최대 매칭(maximum matching, or maximum cardinality matching, MCM)은 가능한 간선 수를 최대로 갖는 매칭으로 매칭 수를 $\nu(G)$ 로 표기한다.

그래프 G_1 에서 $M = \{\{2,3\}\}$ 과 $M = \{\{1,3\}, \{2,4\}\}$ 의 MM이 존재한다. 따라서 MCM은 $\nu(G) = 2$ 인 $M = \{\{1,3\}, \{2,4\}\}$ 가 된다. 따라서 모든 MCM은 MM이 성립하지만 역은 성립하지 않는다.

MCM을 MMEDS(maximum maximal edge dominating set) 또는 MEDS(maximum edge dominating set)라고도 한다.

[정리 1] 일반 그래프의 MCM은 최대 $\nu(G) = \lfloor \frac{n}{2} \rfloor$ 개, $K_{m,n}$ ($m \leq n$) 이분그래프는 최대 $\nu(G) = m$ 이다.

MCM을 구하는 널리 알려진 방법으로 Hopcroft-Karp 알고리즘(HKA)가 있다.^[6] HKA는 이분그래프에 한정하여 수행되며, 너비우선탐색과 깊이우선탐색 법으로 증대경로(augmenting path)를 찾는 방법으로 $O(m \log n)$ 수행 복잡도를 갖고 있다.

그림 3의 G_4 와 G_5 그래프에 대해 HKA를 수행한 결과는 그림 4에 제시되어 있다. HKA는 U 집합에 대해서 임의의 정점부터 매칭 M 에 간선을 계속적으로 추가하면서 증대경로를 확장하는 방법으로 반복 #1에서 미 방문한 정점이 있을 경우 이 정점부터 다시 이 과정을 반복 수행하는 반복 #2를 수행한다.

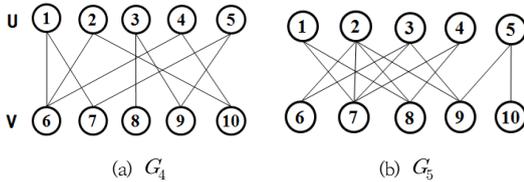


그림 3. G_4 와 G_5 그래프
 Fig. 3. G_4 and G_5 graph

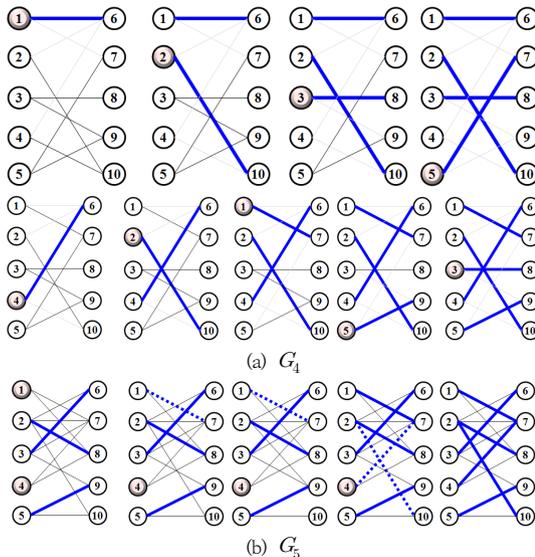


그림 4. G_4 와 G_5 그래프에 대한 HKA
 Fig. 4. HKA for G_4 and G_5 graph 그래프

일반그래프에 대한 MCM을 구하는 알고리즘으로는 $O(\sqrt{nm})$ 복잡도의 Micali-Vazirani 알고리즘(MVA)^[7]이 있으나 매우 복잡하고, 저자는 이 알고리즘을 증명하지 못하였으며, 추후 Vazirani^[12]가 증명하였다.

따라서 3장에서는 최소차수 정점 간 간선을 최대 $\nu(G) = \lfloor n/2 \rfloor = k$ 회 수행하는 $O(n)$ 수행 복잡도 알고리즘을 제안한다.

III. 최소차수 정점 간 간선 선택 알고리즘

이분그래프에 한정하지 않은 어떠한 그래프에서 정점 수 n 에 대해 최대 $\nu(G) = \lfloor n/2 \rfloor$ 의 MCM을 구하려면 가능한 최소차수 정점 간의 간선을 선택해야 한다. 이러한 규칙은 그림 2에서 MM과 MCM을 설명한 부분에서 증명되었으며, 최대독립집합(maximum independent set, MIS)을 구하는 경우와 마찬가지로 논리이다. 따라서 본 장에서는 이러한 단순한 논리에 근거하여 $O(m)$ 의 선행시간 복잡도 알고리즘을 제안하며, 이를 최소차수 정점 간 간선 선택 알고리즘(minimum degree inter-vertex edge selection algorithm, MDVESA)이라 칭한다.

MDVESA는 그림 5에 제시되어 있다. 그래프의 모든 정점에 대한 차수를 구하고, 최소차수 정점 u 와 이 정점에 인접한 정점 $N_G(u)$ 들 중 최소차수 정점 v 를 선정하여 간선 $\{u,v\}$ 를 선정하는 방법으로 일반그래프는 $\lfloor n/2 \rfloor$ 회를, 이분그래프는 $K_{m,n}$ ($m \leq n$)에 대해 m 회 반복 수행한다.

```

• Input :  $G = (V, E)$ ,  $n = |V|, m = |E|$ 
•  $G = (V, E)$ : 정점  $u$ 의 차수  $\text{deg}(u)$  계산 /*  $O(n)$  */
Until  $E = \{\phi\}$  do /*  $O(n/2)$  */
    •  $u = \delta(G)$  정점과  $v = \delta(N_G(u))$  정점의  $\{u, v\}$  간선 결정
    •  $M = \{u, v\}$ ,  $V = V \setminus u, v$ 
    •  $E = E \setminus \{u, w\}, \{v, w\}$ 
    •  $\text{deg}(w) = \text{deg}(w) - 1$ 
End
    
```

그림 5. 최소 차수 정점 간 간선 선택 알고리즘
 Fig. 5. Minimum degree vertex pair edge selection algorithm(MDVESA)

그림 1의 G_1, G_2, G_3 와 그림 3의 G_4, G_5 그래프에 대해 MDVESA를 수행하는 과정은 그림 6에 제시되어 있다. G_4 와 G_5 의 $K_{5,5}$ 이분 그래프에 대해 HKA는 반복 #1에서 $\nu(G) = 5$ 를 구하지 못하고 매칭되지 않은 정점들이

남아 있어 반복 #2에서 이들 정점을 재방문하는 방식으로 수행되었다. 반면에, 제안된 MDVESA는 단순히 $\nu(G)$ 회의 최소 차수 간 정점을 선택하여 해를 구하였다.

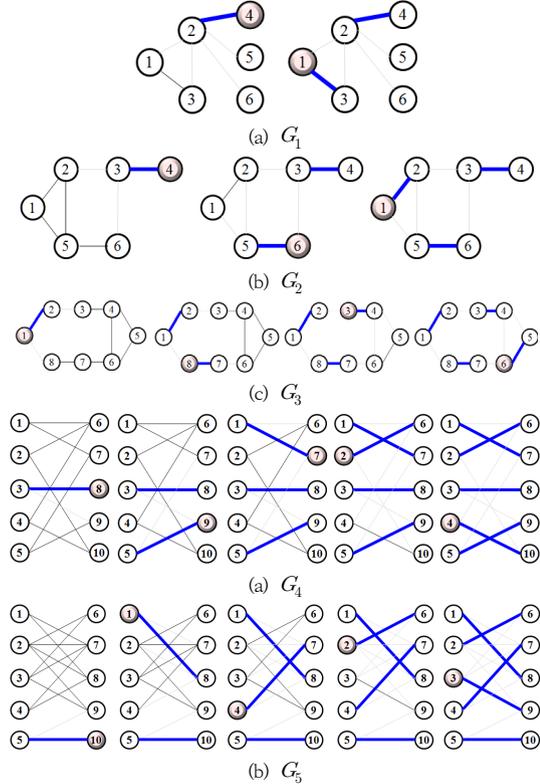


그림 6. G_4 와 G_5 그래프에 대한 MDVESA
Fig. 6. MDVESA for G_4 and G_5 graph 그래프

IV. 알고리즘 적용 및 결과 분석

MDVESA를 그림 7의 다양한 실험 데이터에 적용하여 MCM을 구하여 본 결과는 그림 8에 제시되어 있다.

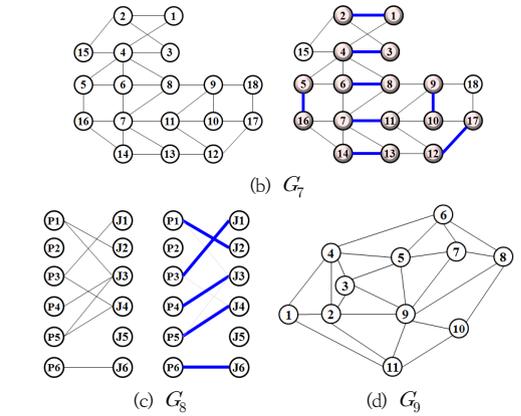
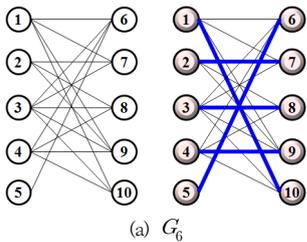
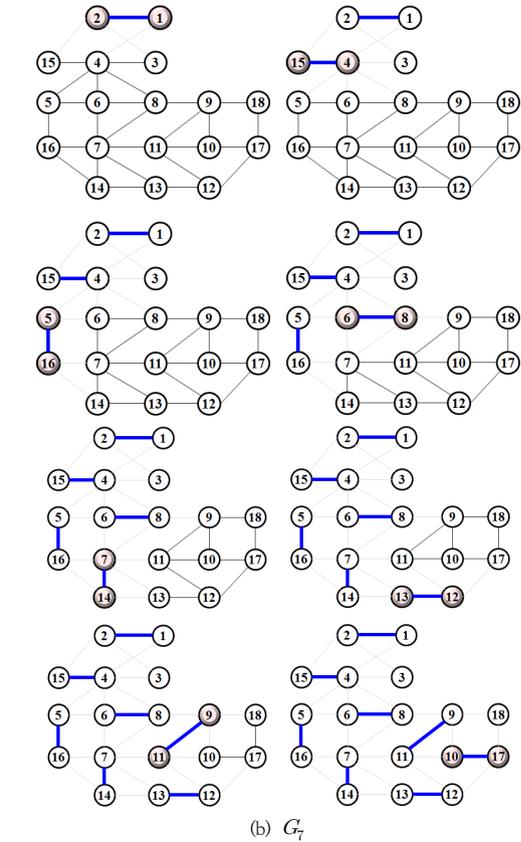
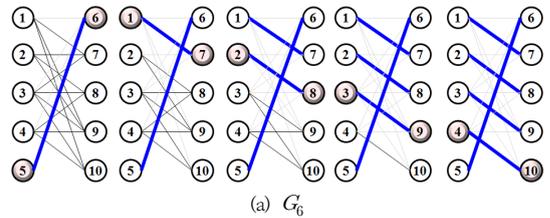


그림 7. 실험 데이터
Fig. 7. Experimental data



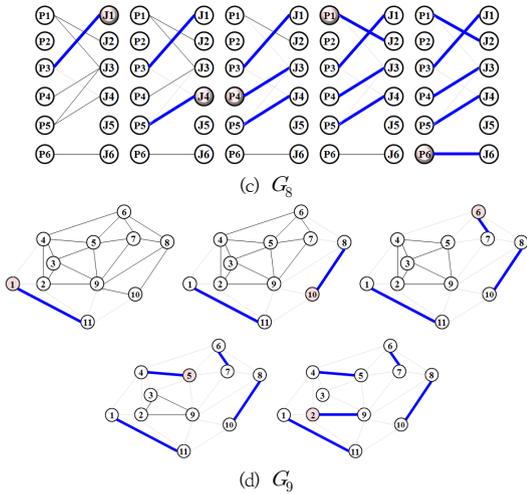


그림 8. 실험 데이터에 대한 MDVESA의 MCM
 Fig. 8. MCM of MDVESA for experimental data

MDVESA를 그림 9의 다양한 추가 실험 데이터에 적용하여 MCM을 구하여 본 결과는 그림 10에 제시되어 있다.

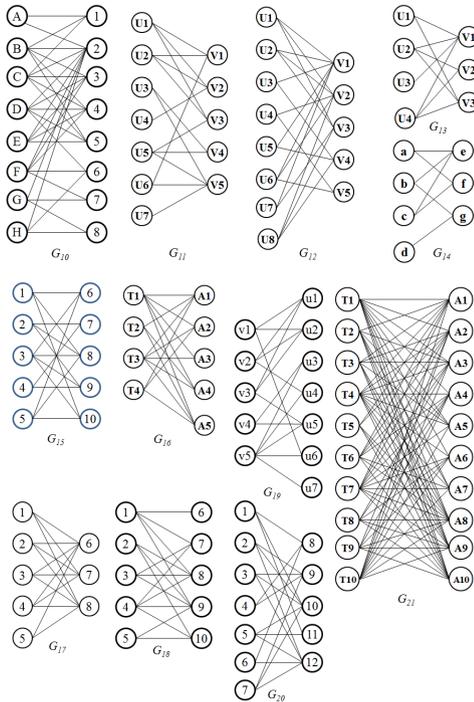


그림 9. 추가 실험 데이터
 Fig. 9. Additional experimental data

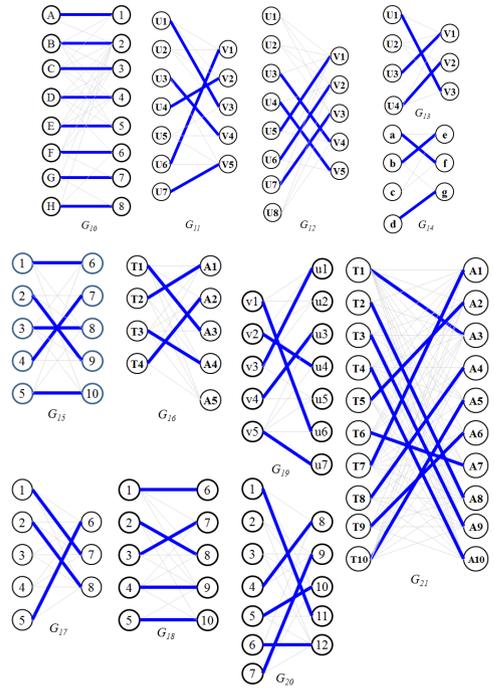


그림 10. 추가 실험 데이터에 대한 MDVESA의 MCM
 Fig. 10. MCM of MDVESA for additional experimental data

표 1. MDVESA 성능
 Table 1. Performance of MDVESA

그래프	종류	정점 수		$\nu(G)$	수행횟수
		$\lfloor n/2 \rfloor$	m		
G_1	일반 $n = 6$	3	-	2	2
G_2	일반 $n = 6$	3	-	3	3
G_3	일반 $n = 8$	4	-	4	4
G_4	이분 $K_{5,5}$	-	5	5	5
G_5	이분 $K_{5,5}$	-	5	5	5
G_6	이분 $K_{5,5}$	-	5	5	5
G_7	일반 $n = 18$	9	-	8	8
G_8	이분 $K_{6,6}$	-	6	5	5
G_9	일반 $n = 11$	5	-	5	5
G_{10}	이분 $K_{8,8}$	-	8	8	8
G_{11}	이분 $K_{7,5}$	-	5	5	5
G_{12}	이분 $K_{8,5}$	-	5	5	5
G_{13}	이분 $K_{4,3}$	-	3	3	3
G_{14}	이분 $K_{4,3}$	-	3	3	3
G_{15}	이분 $K_{5,5}$	-	5	5	5
G_{16}	이분 $K_{4,5}$	-	4	4	4
G_{17}	이분 $K_{5,7}$	-	5	5	5
G_{18}	이분 $K_{5,3}$	-	3	3	3
G_{19}	이분 $K_{5,5}$	-	5	5	5
G_{20}	이분 $K_{7,5}$	-	5	5	5
G_{21}	이분 $K_{10,10}$	-	10	10	10

제안된 MDVESA의 성능을 표 1에 요약되어 있다. 제안된 알고리즘은 최소차수 정점 간 간선을 $\nu(G)$ 회 선택하여 MCM을 구함을 알 수 있다.

V. 결 론

일반 그래프에 대한 MCM은 $O(\sqrt{nm})$ 복잡도 알고리즘이, 이분 그래프에 대해서는 $O(m \log n)$ 의 증대경로 알고리즘이 알려져 있다. 반면에, 본 논문은 주어진 그래프가 일반그래프 또는 이분 그래프와 상관없이 항상 $O(n)$ 복잡도로 최대 매칭(MCM) M 을 구하는 알고리즘을 제안하였다.

제안된 알고리즘은 “최대 매칭을 구하기 위해서는 가능한 많은 정점 쌍의 간선을 선택해야만 한다.”는 기본 원리에 근거하여 최소차수 정점 u 와 $N_G(u)$ 들 중 최소차수 정점 v 간 간선 $\{u, v\}$ 를 $\nu(G) = k$ 회 단순히 선택하는 방법으로 $O(n)$ 수행 복잡도의 간단하면서도 정확한 결과를 얻을 수 있다. 제안된 알고리즘을 일반그래프와 이분그래프의 다양한 실험 데이터들에 적용한 결과 $\nu(G)$ 를 정확하게 구할 수 있음을 보였다.

References

- [1] O. Chaleb, "Introduction to Maximum Matching in Graphs," <http://people.scs.carleton.ca/~maheshwa/courses/5703COMP/16Fall/Matching-Report.pdf>, 2016.
- [2] C. Witzgall and C. T. Zahn, "Modification of Edmonds' Maximum Matching Algorithm" Journal of Research of the National Bureau of Standards, Vol. 69B, No. 1-2, Jan. 1965,
- [3] O. Giel and I. Wegener, "Evolutionary Algorithms and the Maximum Matching Problem," Annual Symposium on Theoretical Aspects of Computer Science, pp. 415-426, Feb. 2003, https://doi.org/10.1007/3-540-36494-3_37
- [4] J. He and X. Tao, "Time Complexity Analysis of an Evolutionary Algorithm for Finding Nearly Maximum Cardinality Matching," Journal of Computer Science and Technology, Vol. 19, No. 4, pp. 450-458, Jul. 2004, <https://doi.org/10.1007/BF02944746>
- [5] V. M Shettar and S. A. Angadi, "A Genetic Algorithm for Graph Matching Using Graph Node Characteristics," Journal of Mathematical Sciences International Research, Vol. 4, No. 2, pp. 347-351, 2015, ISBN:978-93-84124-53-3
- [6] J. E. Hopcroft and R. M. Karp, "An $n^{5/2}$ Algorithm for

Maximum Matchings in Bipartite Graphs," SIAM Journal on Computing, Vol. 2, No. 4, pp. 225-231, Dec. 1973, <https://doi.org/10.1137/0202019>.

- [7] S. Micali and V. V. Vazirani, "An $O(\sqrt{nm})$ Algorithm for Finding Maximum Matching in General Graphs," Proceedings of 21st IEEE Symposium on Foundations of Computer Science, pp. 17-27, Oct. 1980, <https://doi.org/10.1109/SFCS.1980.12>.
- [8] N. Blum, "A New Approach to Maximum Matching in General Graphs," International Colloquium on Automata, Languages, and Programming, pp. 586-597, 1990, <https://doi.org/10.1007/BFb0032060>
- [9] T. Kameda and I. Munro, "A $O(|V||E|)$ Algorithm for Maximum Matching of Graphs," Computing, Vol. 12, No. 1, pp. 91-98, Mar. 1974, <https://doi.org/10.1007/BF02239502>
- [10] Z. Galil, "Efficient Algorithms for Finding Maximum Matching in Graphs," ACM Computing Surveys, Vol. 18, No. 1, pp. 23-38, Mar. 1988, <https://doi.org/10.1145/6462.6502>
- [11] A. Quattoni, X. Carreras, and M. Gallé, "A Maximum Matching Algorithm for Basis Selection in Spectral Learning," Cornell University Library, pp. 1-13, Jun. 2017, arXiv:1706.02857
- [12] V. V. Vazirani, "A Simplification of the MV Matching Algorithm and its Proof," Cornell University Library, pp. 1-32, Aug. 2013, CoRR abs/1210.4594, arXiv:1210.4594

저 자 소 개

이 상 운(정회원)



- 1987년 : 한국항공대학교 항공전자공학과 (학사)
- 1997년 : 경상대학교 컴퓨터과학과 (석사)
- 2001년 : 경상대학교 컴퓨터과학과 (박사)
- 2003년 : 강원도립대학 컴퓨터응용과 전임강사

- 2004년 ~ 2007년 2월 : 국립 원주대학 여성교양과 조교수
- 2007년 3월 ~ 2015년 3월 : 강릉원주대학교 멀티미디어 공학과 부교수
- 2015년 4월 ~ 현재 : 강릉원주대학교 멀티미디어공학과 정교수
- 관심분야 : 소프트웨어 프로젝트 관리, 개발 방법론, 분석과 설계 방법론, 시험 및 품질보증, 소프트웨어 신뢰성, 인공 지능과 빅데이터분석, 최적화 알고리즘
- e-mail : sulee@gwnu.ac.kr