EGERVÁRY RESEARCH GROUP ON COMBINATORIAL OPTIMIZATION



TECHNICAL REPORTS

TR-2002-04. Published by the Egrerváry Research Group, Pázmány P. sétány 1/C, H-1117, Budapest, Hungary. Web site: www.cs.elte.hu/egres. ISSN 1587-4451.

A Gallai-Edmonds-type structure theorem for path-matchings

Bianca Spille and László Szegő

June 2002

1

A Gallai-Edmonds-type structure theorem for path-matchings

Bianca Spille^{*} and László Szegő^{**}

Abstract

As a generalization of matchings, Cunningham and Geelen introduced the notion of path-matchings. We give a structure theorem for path-matchings which generalizes the fundamental Gallai-Edmonds structure theorem for matchings. Our proof is purely combinatorial.

1 Introduction

Cunningham and Geelen in [1] and [2] introduced the notion of path-matchings as a generalization of matchings: Let $G = (V, T_1, T_2; E)$ be an undirected graph and $T_1, T_2 \subseteq V$ disjoint stable sets of G. T_1 and T_2 are called *terminal sets*. We denote $V - (T_1 \cup T_2)$ by R. If $|T_1| = |T_2| =: k$, then a perfect path-matching is a subset $M \subseteq E$ such that the subgraph $G_M = (V, M)$ is a collection of k disjoint paths, all of whose internal nodes are in R, linking the nodes of T_1 to the nodes of T_2 , together with a perfect matching of the nodes of R not in any of the paths. A path-matching with respect to T_1, T_2 is a set M of edges such that every component of the subgraph $G_M = (V, M)$ having at least one edge is a simple path from $T_1 \cup R$ to $T_2 \cup R$, all of whose internal nodes are in R. The one-edge-components in R are called the matching edges of M. The value of a path-matching M is defined to be the number val(M) = |M| + |M'|, where M' denotes the set of the matching edges of M. (That is, the matching edges count twice.) For example, the value of a perfect path-matching is |R| + k. Note that T_1 (and T_2) need not to be stable because edges spanned by T_1 do not play any role here. From now on we do not allow path-matchings having paths in R of length more than 1, that is, any path of a path-matching has at least one end-node in T_1 or T_2 . A path is called a (U, V)-path, if one of its end-nodes is in U and the other in V. For i = 1, 2, a T_i -half-path is a (T_i, R) -path.

^{*}Institute for Mathematical Optimization, University of Magdeburg, Universitätsplatz 2, D-39106 Magdeburg, Germany, e-mail: spille@imo.math.uni-magdeburg.de

Research supported by the European DONET program TMR ERB FMRX-CT98-0202.

^{**}Department of Operations Research, Eötvös University, Pázmány Péter sétány 1/C, Budapest, Hungary H-1117. The author is a member of the Egerváry Research Group (EGRES). e-mail: szego@cs.elte.hu

Research supported by the Hungarian National Foundation for Scientific Research Grant, OTKA T037547, and by FKFP grant no. 0143/2001.

We define a *cut* separating the terminal sets T_1 and T_2 to be a subset $X \subseteq V$ such that there is no path between $T_1 - X$ and $T_2 - X$ in G - X. We denote by $odd_G(X)$ the number of connected components of G - X which are disjoint from $T_1 \cup T_2$ and have an odd number of nodes. Let $Odd_G(X)$ denote the union of these components. Let $Even_G(X)$ denote the union of the components of G - X having an even number of nodes which are disjoint from $T_1 \cup T_2$. For i = 1, 2, let W_i denote the union of components of G - X which are not disjoint from T_i . See Figure 1.

In [4] the following necessary and sufficient condition was proved for the existence of a perfect path-matching and then the following min-max formula was derived for the maximum value of a path-matching.

Theorem 1.1. In $G = (V, T_1, T_2; E)$ there exists a perfect path-matching if and only if $|T_1| = |T_2| = k$ and

$$|X| \ge odd_G(X) + k \qquad for all cuts X.$$

Theorem 1.2. In $G = (V, T_1, T_2; E)$ one has the following formula for the maximum value of a path-matching:

$$\max_{M \text{ path-matching}} val(M) = |R| + \min_{X \text{ cut}} (|X| - odd_G(X)).$$
(1)

Tutte's theorem and the Berge-Tutte-formula are special cases.

A cut X is said to be *tight* if the minimum is attained for it in (1).



Figure 1: A cut X separating T_1 and T_2

A graph G = (V, E) is said to be *factor-critical* if it is connected and each node is missed by a maximum matching.

Lemma 1.3 (Gallai's lemma [5]). If G = (V, E) is factor-critical, then |V| is an odd number and a maximum matching of G has cardinality (|V| - 1)/2.

From Tutte's theorem we obtain

a connected G is factor-critical if and only if $odd_G(Y) \leq |Y|$ for all $Y \subseteq V$, $|Y| \geq 1$. (2) As an easy corollary of Gallai's lemma for a factor-critical graph we have

 $u, v \in V \implies$ there exists a (u, v)-path such that

there exists a perfect matching on the nodes not in the path. (3)

The following theorem plays an important role in Matching Theory.

Theorem 1.4 (The Gallai-Edmonds Structure Theorem [3, 6]). Let G = (V, E) be a graph. Let D denote the set of nodes which are not covered by at least one maximum matching of G. Let A be the set of nodes in V - D adjacent to at least one node in D. Let C = V - A - D. Then:

- The number of covered nodes by a maximum matching in G equals to |V| + |A| c(D), where c(D) denotes the number of components of the graph spanned by D.
- The components of the subgraph induced by D are factor-critical.
- The subgraph induced by C has a perfect matching.
- The bipartite graph obtained from G by deleting C and the edges in A and by contracting each component of D to a single node has the following property: there is a matching covering A after deleting any node obtained by a component of D.
- If M is any maximum matching of G, then $E(D) \cap M$ covers all the nodes except one of any component of D, $E(C) \cap M$ is a perfect matching and M matches all the nodes of A with nodes in distinct components of D.



Figure 2: The Gallai-Edmonds decomposition of a graph G

Here we will prove the following generalization of the Gallai-Edmonds Structure Theorem for path-matchings. Our proof is purely combinatorial and is an extension of the proof of Theorem 1.2 in [4]. The careful investigation of the augmenting path algorithm of Spille and Weismantel [8, 9] for path-matchings gives an algorithmic proof but for the sake of brevity here we omit the details.

Define

$$\nu(G = (V, T_1, T_2; E)) := \max_{M \text{ path-matching in } G} val(M).$$

Theorem 1.5 (Structure Theorem for Path-Matchings). Let $G = (V, T_1, T_2; E)$ be a graph. Define the following sets.

$$F := \{ v \in R : \nu(G - v) = \nu(G) \},\$$

$$F_1 := \{ v \in R : \nu(G' = (V, T_1, T_2 + v; E)) = \nu(G) \},\$$

$$F_2 := \{ v \in R : \nu(G'' = (V, T_1 + v, T_2; E)) = \nu(G) \},\$$

$$H_1 := \{ v \in T_1 : \nu(G - v) = \nu(G) \},\$$

$$D_1 := F \cup F_1 \cup H_1,\$$

$$A_1 := \{ v \in V - D_1 : \exists u \in D_1 \text{ such that } uv \in E \} \cup (T_1 - D_1)\$$

$$C_1 := V - A_1 - D_1.$$

Then:

- (S1) A_1 is a cut and $\nu(G) = |R| + |A_1| odd_G(A_1)$ (that is, A_1 is a tight cut).
- (S2) The components of the subgraph induced by D_1 and disjoint from T_1 are factorcritical.
- (S3) F is the union of some components of D_1 which are disjoint from T_1 .
- (S4) $F_1 \cap F_2 \subseteq F$ and $F_1 \cap F_2$ is the union of some components of D_1 disjoint from T_1 .
- (S5) The components of the subgraph induced by C_1 which are disjoint from T_1 and T_2 have a perfect matching.
- (S6) For any component K of F, there is a maximum path-matching M for which there is no edge of M coming out of K.
- (S7) If M is any maximum path-matching of G, then $val(E(K) \cap M) = |K| 1$ for any component K of $F \cup F_1$ which is disjoint from T_1 .
- (S8) If M is any maximum path-matching of G, then $val(E(C_1) \cap M) = |C_1|$.
- (S9) If M is any maximum path-matching of G, then any component K of D_1 is either traversed by one path P of M and $K \cap P$ is connected, or there is exactly one matching edge with one end-node in K and the other in X, or there is no edge of M coming out of K; and there is no edge of M spanned by X, and there is no edge of M coming out of any even component of G - X which is disjoint from T_1 and T_2 .

We may define D_2, A_2, C_2 similarly, that is, surprisingly there are two kinds of structure theorems for path-matchings.

The special case of $T_1 = T_2 = \emptyset$ gives the original Gallai-Edmonds structure theorem: $F_1 - F = F_2 - F = H_1 = \emptyset$, $D = F = D_1$, $A = A_1$, $C = C_1$.

The above sets F, F_1 , F_2 , and H_1 can be interpreted as follows:

F is the set of nodes $v \in R$ for which there is a maximum path-matching M not

covering v. $F_i - F$ is the set of nodes $v \in R$ which are not in F and there is a maximum path-matching M so that v is an end-node of a T_i -half-path of M (i = 1, 2). H_1 is the set of nodes $v \in T_1$ for which there is a maximum path-matching M not covering v.

 $F_1 \cap F_2 \subseteq F$ means that if a node v is an end-node of a T_i -half-path for a maximum path-matching K_i for i = 1, 2, then there is a maximum path-matching K not covering v.

For G = (V, E) and $K \subseteq V$, define $E[K] := \{uv \in E : u, v \in K\}$ and G[K] := (K, E[K]).

2 Proofs

2.1 (Optimality Criteria). Let M be a path-matching and X a cut in G. M is a maximum path-matching and X is a tight cut if and only if the following statements hold:

- (O1) M induces a perfect matching on $Even_G(X)$, $val(E[Even_G(X)] \cap M) = |Even_G(X)|.$
- (O2) For any component K of $Odd_G(X)$, M induces a matching and an even path (possibly \emptyset) on K covering all (but possibly one) nodes of K, $val(E[K] \cap M) = |K| 1$.
- (O3) For i = 1, 2, M induces T_i -half-paths and matching edges on W_i covering all the nodes of $W_i T_i$.
- (O4) For any node $v \in X$, v is either covered by a matching edge of M, by a (T_1, T_2) path of M, or by a T_i -half-path of M but v is not the R-end-node (i = 1, 2). Minduces no edge on X.
- (O5) For any R-end-node v of a T_i -half-path of $M, v \in Odd_G(X) \cup W_i$ (i = 1, 2). For any $v \in R$ not covered by $M, v \in Odd_G(X)$.

Proof. If (O1)–(O5) hold, then

$$val(M) = |Even_G(X)| + |Odd_G(X)| - odd_G(X) + |X| + |W_1 - T_1| + |W_2 - T_2|$$

= |R| + |X| - odd_G(X),

which proves that M is maximum and X is tight.

If M is a maximum path-matching and X is a tight cut, then let P_1, P_2, \ldots, P_n denote the (T_1, T_2) -paths of M, and let $P'_1, P'_2, \ldots, P'_{n_1}$ denote the T_1 -half-paths traversing X and $P''_1, P''_2, \ldots, P''_{n_2}$ denote the T_2 -half-paths traversing X. For a path P_i (P'_i, P''_i) , let t_i (t'_i, t''_i) respectively) denote the number of components of $Odd_G(X)$ which are traversed by P_i (P'_i, P''_i) respectively). Orient the edges of these paths from T_1 to T_2 . We have

$$\alpha \le \sum_{i=1}^{n} t_i + \sum_{i=1}^{n_1} t'_i + \sum_{i=1}^{n_2} t''_i, \tag{4}$$

where α denotes the number of components of $Odd_G(X)$ which are traversed by some path P_i, P'_i, P''_i . Let β denote the number of components K of $Odd_G(X)$ for which a matching edge of M has one end-node in K and the other in X, and no path of M traverses K. Let $\gamma := odd_G(X) - \alpha - \beta$, i.e., γ is the number of components of $Odd_G(X)$ not traversed by any edge of M. Since any of the paths P_i has a first node in X and for any of the paths P_i, P'_i before traversing a component of $Odd_G(X)$ there is a node in X, and for any of the paths P_i, P''_i after traversing a component of $Odd_G(X)$ there is a node in X, we have

$$n + odd_G(X) - \gamma = n + \alpha + \beta \le \sum_{i=1}^n (t_i + 1) + \sum_{i=1}^{n_1} t'_i + \sum_{i=1}^{n_2} t''_i + \beta \le |X|, \quad (5)$$

since we determined distinct nodes of X. Hence, $n - \gamma \leq |X| - odd_G(X)$. Since M is maximum and X is tight, we obtain

$$val(M) = |R| + |X| - odd_G(X) \ge |R| - \gamma + n$$

The value of M is equal to the number of nodes in R covered by M plus the number of (T_1, T_2) -paths of M (which is n). Hence, the number of nodes in R not covered by M is less than or equal to γ . Since any component of $Odd_G(X)$ not traversed by any edge of M contains at least one node not covered by M, equality holds through. Hence, we have equality in (5) and (4). We obtain (O1)–(O5).



Figure 3: A maximum path-matching M and a tight cut X

Proof of Theorem 1.5. Let X be a tight cut for which the union of components of G - X which are not disjoint from T_1 and the odd components which are disjoint from $T_1 \cup T_2$ is minimal, furthermore $X \cap T_1$ is maximal. Define $D_X := W_1 \cup Odd_G(X)$.

Claim 2.2. Each component of $G[D_X]$ disjoint from T_1 is factor-critical.

Proof. Let K be a component of $G[D_X]$ disjoint from T_1 . If K has an even number of nodes, then X + v is a tight cut and $D_{X+v} \subseteq D_X - v$ for $v \in K$, contradicting the choice of X. Hence, K has an odd number of nodes. Let $Y \subseteq K$ be a subset with $odd_{G[K]}(Y) > |Y|$. Since

$$|X \cup Y| - odd_G(X \cup Y) = |X| - odd_G(X) + |Y| - odd_{G[K]}(Y) + 1 \le |X| - odd_G(X),$$

 $X \cup Y$ is a tight cut and $D_{X \cup Y} \subseteq D_X - Y$. The choice of X implies $Y = \emptyset$. Now (2) implies that K is factor-critical.

We will prove that $D_1 = D_X$, $A_1 = X$, and $C_1 = V - (X \cup D_X)$.

Without loss of generality, $X \neq T_1$ or $Odd_G(T_1) \neq \emptyset$. Let us contract each component of $Odd_G(X)$ to a node. Let Q denote the set of new nodes and let G_Q denote the graph obtained this way. Notice that $|Q| = odd_G(X)$.

Claim 2.3. If G_Q has a path-matching of value k, then G has a path-matching of value $k + |Odd_G(X)| - odd_G(X)$.

Proof. Let M_Q denote the path-matching of G_Q with value k. Let M denote the set of edges of G corresponding to M_Q . We claim that M can be completed in G to be a path-matching with the desired value. To this end, let K denote a component of $Odd_G(X)$, and let q denote its corresponding node in G_Q . By Claim 2.2, K is factor-critical.

If M_Q covers q by a matching edge, then M covers one node, say v, of K, and by Gallai's lemma there is a perfect matching on K - v. If M_Q covers q by a path, then M covers either one node v of K or two distinct nodes, say u and v, of K. In the first case, Gallai's lemma applies again, while in the second one, by (3), there is a path P in K connecting u and v and a perfect matching on K - V(P), where V(P) denotes the nodes of P. If M_Q does not cover q, then Gallai's lemma applies again.



Figure 4: G_l and G_r

Claim 2.4. Let $V_l := Q \cup W_1 \cup (X - T_1), T_1^l := (T_1 - X) \cup Q$, and $T_2^l := X - T_1$. Then $X - T_1$ is the unique tight cut in G_l , i.e.,

$$\nu(G_l = (V_l, T_1^l, T_2^l; E_l)) = |R_l| + |X - T_1|$$

and for any cut $Y \neq X - T_1$ in G_l , $|Y| - odd_{G_l}(Y) \ge |X - T_1| + 1$.

Proof. $X - T_1 = T_2^l$ is a cut in G_l with $odd_{G_l}(X - T_1) = 0$. Let Y be a tight cut in G_l , then $|Y| - odd_{G_l}(Y) \leq |X - T_1|$. Denote $Z := (Y - Q) \cup (T_1 \cap X)$. Since X is a cut in G and Y is a cut in G_l , Z is a cut in G. We have $odd_G(Z) \geq odd_{G_l}(Y) + |Q - Y|$ and $D_Z \subseteq D_X$. Hence,

$$\begin{aligned} |Z| - odd_G(Z) &\leq (|Y - Q| + |T_1 \cap X|) - (odd_{G_l}(Y) + |Q - Y|) \\ &= |Y| - odd_{G_l}(Y) + |T_1 \cap X| - |Q| \\ &\leq |X - T_1| + |T_1 \cap X| - odd_G(X) \\ &= |X| - odd_G(X). \end{aligned}$$

Since X is tight, Z is a tight cut. By the choice of X, $D_Z = D_X$ and $|X \cap T_1| \ge |Z \cap T_1|$. This implies $Y = X - T_1$.



Figure 5: A tight cut in G_l

Analogously, we obtain

Claim 2.5. Let $V_r := Q \cup W_2 \cup (X - T_2), T_1^r := (X - T_2), and T_2^r := (T_2 - X) \cup Q.$ Then $X - T_2$ is a tight cut, i.e.,

$$\nu(G_r = (V_r, T_1^r, T_2^r; E_r)) = |R_r| + |X - T_2|.$$

Claim 2.6. $D_X = D_1$.

Proof. (O5) implies $D_1 \subseteq D_X$. It remains to prove $D_X \subseteq D_1$. Let $v \in D_X = W_1 \cup Q$. Let Y be a cut in $G' = G_l - v$. Then Y + v is a cut in G_l and Claim 2.4 implies

$$|Y| + 1 - odd_{G'}(Y) = |Y + v| - odd_{G_l}(Y + v) \ge |X - T_1| + 1$$

and hence,

$$|R_l| + \min_{Y \text{ cut in } G'} (|Y| - odd_{G'}(Y)) \ge |R_l| + |X - T_1| = \nu(G_l).$$

If $v \in (T_1 - X) \cup Q = T_1^l$ then $R' = R_l$ implying that $\nu(G') = \nu(G_l)$ and hence, there exists a maximum path-matching M_l in G_l not covering v. If $v \in W_1 - T_1 = R_l$, then $|R'| = |R_l| - 1$ implying that $\nu(G') \ge \nu(G_l) - 1$ and hence, there exists a maximum path-matching M_l in G_l such that v is not covered by M_l or v is an end-node of a T_1^l -half-path of M_l . By Claim 2.5, there is a path-matching M_r of G_r not covering vwith value $|R_r| + |X - T_2|$. By (O1), there is a perfect matching M_E on $Even_G(X)$. Now $M' := M_l \cup M_r \cup M_E$ is a nearly path-matching of G_Q , where a nearly pathmatching is the disjoint union of a path-matching and some even cycles lying entirely in R. Its value is the value of the path-matching plus the number of edges in even cycles, hence,

$$val(M') = |R_l| + |X - T_1| + |R_r| + |X - T_2| + |Even_G(X)| = |R| + |X| - |Odd_G(X)|.$$

Moreover, v is not covered by any edge of M' or v is an end-node of a T_1 -half-path of M'. Transforming the even cycles of M' into the union of matching edges, we obtain a path-matching M^* of G_Q of the same value.

By Claim 2.3, G has a path-matching M with value

$$val(M) = val(M') + |Odd_G(X)| - odd_G(X) = |R| + |X| - odd_G(X)$$

and v is not covered by M or v is an end-node of a T_1 -half-path of M. By Theorem 1.2, M is a maximum path-matching. Consequently, $v \in D_1$.

Next we show $A_1 = X$. By definition, $A_1 = (\text{neighbors of } D_1 - D_1) \cup (T_1 - D_1)$. Since $D_1 = W_1 \cup Odd_G(X)$, it follows $T_1 \cap X \subseteq A_1 \subseteq X$. Let $v \in X - T_1$. By (O4), v has a neighbor w in R - X. By (O1), $w \notin Even_G(X)$ and by (O3), $w \notin W_i - T_i$ (i = 1, 2). Hence, $w \in Odd_G(X) \subseteq D_1$. Consequently, $v \in A_1$. This proves $A_1 = X$ and (S1) follows.

Because of $D_X = D_1$, (S2) is a corollary of Claim 2.2.

Now we prove (S3). (O5) implies $F \subseteq Odd_G(X)$. Let K be a component of $Odd_G(X)$ such that $K \cap F \neq \emptyset$. Let $v \in K \cap F$. Then there exists a maximum pathmatching M not covering v. Since K is factor-critical, for any node $w \in K$ there is a maximum matching M_w in K not covering w. Hence, $M - M[K] \cup M_w$ is a maximum path-matching not covering w, thus, $w \in F$. This implies $K \subseteq F$. Consequently, F is the union of some components of $Odd_G(X)$, i.e., (S3) holds.

Next we show (S4). Let $v \in F_1 \cap F_2$. (O5) implies $v \in Q$. Hence, there exists a maximum path-matching M_l in G_l such that v is not covered by M_l and there exists a maximum path-matching M_r in G_r such that v is not covered by M_r . The same construction as in the proof of $D_X = D_1$ leads to a maximum path-matching M in G not covering v, i.e., $v \in F$. Consequently, $F_1 \cap F_2 \subseteq F$. Similar arguments as for (S3) show that $F_1 \cap F_2$ is the union of some components of D_1 which are disjoint from T_1 .

(S5) follows from $C_1 = V - (X \cup D_X) = Even_G(X) \cup W_2$ and (O1).

(S6), (S7), (S8), and (S9) are direct corollaries of the Optimality Criteria. \Box **Remark.** In [7] by Lovász and Plummer the following structure theorem was given for bipartite graphs. It easily follows from Theorem 1.5. **Theorem 2.7.** Let $G = (U_1, U_2; E)$ be a bipartite graph and for i = 1, 2, let $A_i := A \cap U_i$, $C_i := C \cap U_i$, and $D_i := D \cap U_i$, where A, C, and D are the three sets of the Gallai-Edmonds structure theorem for G. Then

- $D = D_1 \cup D_2$ does not induce any edge of G,
- the subgraph $G[C_1 \cup C_2]$ has a perfect matching and hence, $|C_1| = |C_2|$,
- $N_G(D_1) = A_2$ and $N_G(D_2) = A_1$,
- every maximum matching of G consists of a perfect matching of G[C₁ ∪ C₂], a matching of A₁ into D₂ and a matching of A₂ into D₁,
- if T is any minimum node-cover (i.e. cut) for G,

$$A_1 \cup A_2 \subseteq T \subseteq A_1 \cup A_2 \cup C_1 \cup C_2,$$

- $C_1 \cup A_1 \cup A_2$ and $C_2 \cup A_1 \cup A_2$ are minimum node-covers (i.e. cuts). Consequently, $A_1 \cup A_2$ is the intersection of all minimum node-covers (i.e. cuts), and
- the subgraphs induced by $A_1 \cup D_2$ and $A_2 \cup D_1$ have positive surplus when viewed from A_1 and A_2 respectively.

References

- W. H. Cunningham and J. F. Geelen, *The Optimal Path-Matching Problem*, Proceedings of thirty-seventh Symposium on the Foundations of Computing, IEEE Computer Society Press (1996), 78–85.
- [2] W. H. Cunningham and J. F. Geelen, The Optimal Path-Matching Problem, Combinatorica, 17/3 (1997), 315–336.
- J.R. Edmonds, Maximum matching and a polyhedron with 0,1-vertices, J. Res. Nat. Bur. Standards Sect. B (1968), 125–130.
- [4] A. Frank and L. Szegő, A note on the path-matching formula, Journal of Graph Theory, to appear.
- [5] T. Gallai, Neuer Beweis eines Tutte'schen Satzes, Magyar Tud. Akad. Mat. Kutató Int. Közl., 8 (1963), 135–139.
- [6] T. Gallai, Maximale Systeme unabhänginger Kanten, Magyar Tud. Akad. Mat. Kutató Int. Közl., 9 (1965), 401–413.
- [7] L. Lovász, M.D. Plummer, Matching Theory, Akadémiai Kiadó, Budapest, 1986.
- [8] B. Spille and R. Weismantel, A combinatorial algorithm for the independent pathmatching problem, manuscript (2001).

[9] B. Spille and R. Weismantel, A generalization of Edmonds' matching and matroid intersection algorithms, Proceedings of the Ninth International Conference on Integer Programming and Combinatorial Optimization, Lecture Notes in Computer Science 2337, Springer (2002), 9–20.