2
Preliminaries

As stated in the introduction, given a tree T rooted at node r, an assignment A and a weight function w, the cost of A under the weights w is given by $\mathrm{EP}(T, A, w) = \sum_{u \in T} d(r, u, T + A) w(u)$. (We omit the weight function when it is clearly understood from the context.) Furthermore, we extend this definition to subtrees of T: for any subtree T' of T, $\mathrm{EP}(T, A)_{T'} = \sum_{u \in T'} d(r, u, T + A) w(u)$ indicates the expected cost of reaching nodes of T' in the enhanced tree $T + A$. Also, $\mathrm{OPT}_k(T, w)$ denotes the cost of the optimal k-assignment for T with respect to the weights w (henceforth we use $\mathrm{OPT}(T, w)$ as a shorthand for $\mathrm{OPT}_1(T, w)$).

In addition, for any subset U of nodes of T, $w(U)$ denotes the sum of the weights of the elements of U, namely $w(U) = \sum_{u \in U} w(u)$. For each node u of T we define T_u as the subtree of T composed by all descendants[1] of u. For any tree T, we use $r(T)$ to denote the root of T. Also, for every tree T we use $\mathit{height}(T)$ to denote the height of T, that is, the length (in number of arcs) of the largest path from $r(T)$ to a node $u \in T$. Similarly, for every enhanced tree $T + A$, $\mathit{height}(T + A)$ is defined as the length of the largest *user path* in $T + A$ from $r(T)$ to a node $u \in T$. Finally, we extend the set difference operation to trees: given trees $T^1 = (V^1, E^1)$ and $T^2 = (V^2, E^2)$, $T^1 - T^2$ is the forest of T^1 induced by the nodes $V^1 - V^2$.

A concept that is helpful during the analysis of the results is that of a *non-crossing* assignment. Two hotlinks (u, a) and (v, b) for T are crossing if u is an ancestor of v, v is an ancestor of a and a is an ancestor of b (Figure 1.1.b). An assignment is said to be *non-crossing* if it does not contain crossing hotlinks. Using the definition of the greedy model, it is not difficult to see that any crossing assignment can be transformed into a non-crossing one via removal of some hotlinks, and that these removals do not affect the expected path length.

The next proposition is a direct implication of the definition of a valid hotlink assignment.

[1]By definition both the set of ancestors and the set of descendants of a node u include u. In order to exclude u, we refer to proper ancestors of proper descendants.

Proposition 1 *Consider a tree T and an assignment A for it. Let u and v be nodes in T such that $v \in T_u$. Let T' be a subtree of T that contains both u and v. Then, the user path from u to v in $T + A$ equals to the user path from u to v in $T' + A$, and consequently in $T' + A'$, where A' is the set of hotlinks of A with both endpoints in T'.*

Another related proposition, which can be easily proved by induction, is the following:

Proposition 2 *Consider a tree T and an assignment A for it. Let u and v be nodes of T and let P be the path in $T + A$ from u to v. Also consider an assignment A' such that for each $u' \in P$ and for each ancestor v' of v the hotlink (u', v') belongs to A' if and only if it also belong to A. Then the path from u to v is the same in $T + A$ and $T + A'$.*

Now we state two important structural lemmas that allow us to perform transformations on hotlink assignments without increasing much the expected user path length (proofs in the appendix).

Lemma 1 (Multiple Removal Lemma) *Consider a tree T rooted at node r and a weight function w. Let A be an assignment for T with at most g hotlinks leaving r and at most one hotlink leaving every other node. Then, there is an assignment A' with at most one hotlink per node such that $EP(T, A') \leq EP(T, A) + (g - 1)w(T)$.*

Lemma 2 *Consider a tree T and a weight function w. Let T' be a subtree of T. If $v \in T$ is an ancestor of $r(T')$, then $\sum_{u \in T'} d(v, u, T + A)w(u) \geq OPT_g(T', w)$ for any g-assignment A.*

Corollary 1 (Supermodularity) *Consider a tree T and a weight function w. Let $\{T^1, T^2, \ldots, T^k\}$ be pairwise disjoint subtrees of T. Then $OPT_g(T, w) \geq \sum_{i=1}^{k} OPT_g(T^i, w)$.*

Proof: Let A^* be an optimal g-assignment for T. As the trees $\{T^i\}$ are pairwise disjoint, the non-negativity of both $d(.)$ and $w(.)$ implies that:

$$\begin{aligned}
\mathrm{OPT}_g(T, w) &= \sum_{u \in T} d(r(T), u, T + A^*)w(u) \geq \sum_{i=1}^{k} \sum_{u \in T^i} d(r(T), u, T + A^*)w(u) \\
&\geq \sum_{i=1}^{k} \mathrm{OPT}_g(T^i, w)
\end{aligned}$$

where the last inequality follows from Lemma 2. ■

The following lemma generalizes the well known fact that every tree U has a node, say u, such that all trees in the forest $U \setminus u$ have at most $|U|/2$ nodes and can be proved in a similar way.

Lemma 3 *Consider a tree U, a weight function w and a constant α. Then, there is a partition of U into subtrees such that each, except possibly the one containing $r(U)$, has weight with respect to w greater than α. In addition, for every tree U^i in the partition, each of the subtrees rooted at the children of $r(U^i)$ have weight not greater than α.*