13: Metric TSP

בעיית הסוכן הנוסע המטרית

הקבוצה $[n]$ מייצגת ערים. יש מטריצת עלויות סימטרית: $C \coloneqq (c_{ij} \geq 0)$. $c_{ii} = 0$ לכל i.

הבעיה נקראת ה-*TSP* ***metric*** בגלל ההנחה שהמטריצה מקיימת את אי"ש המשולש. אפשר לראות את הבעיה בתור גרף ממושקל מלא.

המטרה היא למזער את העלויות (או המשקל) של מעגל המילטון בגרף.

אבחנה מרכזית: אם ניקח את המעגל של ה-*TSP* (להלן *tour*, סיור), נקבל עץ פורש.

ולכן, המשקל של הסיור האופטימלי הוא לפחות המשקל של העץ פורש מינימום.

Nearest Extension Algorithm

אתחול:

1. ניקח את $i, j \in [n]$, שני הקודקודים הכי קרובים.
2. נגדיר את הסיור להיות הגרף המכוון: $T \coloneqq i \longrightarrow j$.

צעד *extension*. לא נוסיף קודקודים לפני i או אחרי j, אלא רק ביניהם:

3. נניח ש $V(T) \neq [n]$.

4. תהי $y \in [n] \setminus V(T)$ הנקודה שהכי קרובה למסלול, ו-x הנקודה שהגדירה את המרחק למסלול (הכי קרובה ל-y):
$$(x, y) \coloneqq argmin\{c_{xy} : \ x \in V(T), \ \ y \notin V(T)\}$$
אפשר לחשוב על זה כך: הגדרנו חתך שבו צד אחד זה $V(T)$, וכל פעם ניקח את הקודקוד שמחובר ע"י הצלע הקלה ביותר שחוצה את החתך.

5. "נבלע" (*absorb*) את y בצורה הבאה:
נסתכל על הקודקוד הבא במסלול אחרי x (נגיד z), ונעשה עיקוף: במקום $x \to z$, נעשה $x \to y \to z$.
נשים לב: לכאורה אין לנו שליטה על המשקל שמוסיפים עם yz. אבל, נוכל לרסן את זה עם אי"ש המשולש.
   a. אם $x = j$, ניקח את הקודקוד שלפניו במקום אחריו.

סיום:

6. כאשר $V(T) = [n]$, יש לנו סיור T שמתחיל ב-i ועובר בכל הקודקודים ומסתיים ב-j.
נזכור, שב-*extension* הראשון זרקנו את הצלע ij. אז נחזיר את הצלע ji.

טענה: האלגוריתם הזה נותן 2-קירוב עבור בעיית *Metric-TSP*.

הוכחה: יהי G הגרף הממושקל המלא, ונסמן F' את קבוצת הצלעות שהחמדן בחר (הצלע שהייתה מינימלית בחתך בכל שלב).

נגדיר $F \coloneqq F' \cup \{ij\}$. נשים לב ש $([n], F)$ הוא עפ"מ (לפי תהליך הבנייה, אלגוריתם פרים – *prim*).

כלומר, כל צלע של F' היא צלע מינימלית של חתך. הגודל של F' הוא $|F'| = n - 2$, אז $|F| = n - 1$.

וקיימת הרצה של פרים לפי הסדר של הצלעות שהאלגוריתם בחר, שתיתן את F בתור עפ"מ.

ולכן, אם נסמן OPT את העלות של הסיור האופטימלי, נקבל:

$$OPT \geq \sum_{f \in F} c(f)$$

כי אנחנו יודעים שהמשקל של הסיור האופטימלי הוא לפחות המשקל של העץ פורש המינימום.

יהי $T + ij$ הסיור שמוחזר מהאלגוריתם. הסיור הזה התקבל ע"י סדרת מעקפים.

במעקף $u \to v \implies u \to w \to v$, העלות עלתה ב:

$$+ c_{uw} + c_{wv} - c_{uv} \geq 0$$

בגלל אי"ש המשולש, זה לפחות 0. ואנחנו יודעים גם:

$$c_{wv} \leq c_{wu} + c_{uv}$$

אז נוכל לרשום:

$$increase := c_{uw} + c_{wv} - c_{uv} \le c_{uw} + \underbrace{c_{wu} + c_{uv}}_{\le c_{wv}} - c_{uv} = 2c_{wu}$$

וניזכר ש: $uw \in F'$, ושהיא הייתה בחירה חמדנית. כלומר העלות של כל הסיור היא: לכל היותר, העלות של כל המעקפים ב- F', ועוד c_{ij}:

$$cost\ of\colon\ T + ij \le [cost\ of\ all\ bypasses\ in\ F'] + c_{ij} \le c_{ij} + \sum_{f \in F'} 2 \cdot c(f) \le$$

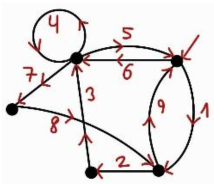
נכפיל את c_{ij} ב-2 כדי להכניס אותו גם לסכום.

כזכור, הצלעות של F מהוות MST. ואמרנו שה-MST האופטימלי הוא לכל היותר ה-$tour$ האופטימלי (נסמן OPT), כלומר נוכל לרשום:

$$\le 2 \cdot \underbrace{\sum_{f \in F} c(f)}_{MST} \le 2 \cdot OPT$$

Doubling Trees Algorithm

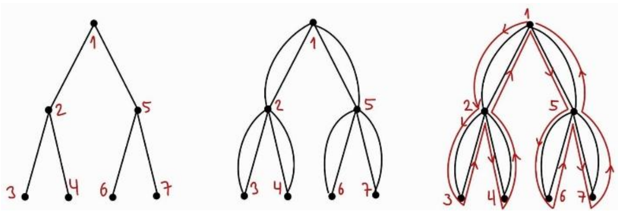
תזכורת: **מעגל אוילר** ($Euler\ tour$) על מולטיגרף הוא הילוך סגור שעובר על כל צלע בדיוק פעם אחת:



גרף שיש בו מעגל כזה נקרא אוילרי, או אוילריאני. מציאת מסלול אוילר היא בעיה ב-P.

<u>משפט</u>: מולטיגרף קשיר יהיה אוילרי **אמ"מ** לכל הקודקודים יש דרגה זוגית. אינטואיטיבית, כי צריך לצאת ולהיכנס מכל קודקוד פעם אחת.

מה הקשר בין בעיית TSP למעגל אוילר? <u>אבחנה</u>: אם T הוא עץ, אז הכפלת כל צלע נותנת לנו גרף אוילרי:



כי אין מעגלים, והגרף קשיר. אז אם מכפילים כל צלע, לכל קודקוד יש דרגה זוגית. זה נותן לנו את האלגוריתם הבא למציאת סיור:

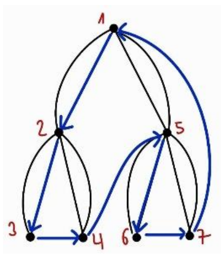
1. ניקח את T, עפ"מ על G.
2. נכפיל כל צלע ב-T (עם המשקלים), ויהי $\mathcal{T}$ המולטיגרף המתקבל.
3. נמצא מעגל אוילר ε ב- $\mathcal{T}$.
4. נבנה סיור TSP מתוך ε ע"י $shortcutting$.

ביצוע $shortcutting$ (קיצור דרך):

ניקח את הקודקודים לפי הסדר שהם מופיעים ב-ε בפעם הראשונה, והקודקוד הראשון חוזר על עצמו. בדוגמה לעיל יש לנו:

$$\varepsilon = (\bar{1}, \bar{2}, \bar{3}, 2, \bar{4}, 2, 1, \bar{5}, \bar{6}, 5, \bar{7}, 5, 1) \implies T := (1,2,3,4,5,6,7,1)$$

ונקבל את הסיור:



טענה: האלגוריתם $doubling\ trees$ הוא 2-מקרב.

הוכחה: נגדיר: OPT את העלות של סיור TSP אופטימלי, L את הסיור המתקבל מהאלגוריתם, T את ה-MST בתהליך ההרצה, ו-$\mathcal{T}$ את העץ ה"מוכפל". מתקיים:

$$c(\mathcal{T}) = 2 \cdot c(T) \le 2 \cdot OPT$$

כי $\mathcal{T}$ זה פשוט T עם כל הצלעות פעמיים, כולל המשקלים. ואנחנו יודעים שמשקל ה-MST חסום בעלות הסיור האופטימלי.

בגלל שהמשקלים מקיימים את אי"ש המשולש, כל קיצור דרך לא מעלה את העלות. אז המשקל הסופי של הסיור הוא לכל היותר מה שהתחלנו איתו. התחלנו עם $\mathcal{T} \le 2 \cdot OPT$, אז המשקל הסופי הוא גם לכל היותר $2 \cdot OPT$.

Christofides Algorithm

בניית גרף אוילר ע"י הכפלת כל הצלעות של MST זה תהליך בזבזני. הקודקודים הבעייתיים הם רק אלה עם דרגה אי"ז.

אבחנה: בכל גרף, מספר הקודקודים עם דרגה אי"ז הוא תמיד זוגי.

אז אם G' גרף קשיר, ו-M הוא שידוך מושלם של <u>קודקודי G' שבעלי דרגה אי"ז</u> (עם צלעות שאולי לא ב- $E(G')$! יכול להיות שנוסיף צלעות).

אז זה בעצם הוסיף 1 לכל קודקוד אי"ז, ועכשיו לכל הקודקודים יש דרגה זוגית, והגרף $G' + M$ הוא אוילרי.

לפי זה, נקבל אלגוריתם:

1. יהי T עפ"מ של G. (ה-T הזה יהיה ה- G' מהטענה לעיל).
2. תהי O קבוצת הקודקודים ב-T שבעלי דרגה אי"ז.
3. יהי M שידוך מושלם בעל מחיר מינימום בגרף $G[O]$ (הגרף המלא על הקודקודים ב-O). מציאת ה-M המינימום היא בעיה ב-P.
4. נוסיף את M ל-T כדי לקבל את $\mathcal{T}$, שהוא אוילרי לפי הטענה.
5. נמצא סיור אוילר ε ב- $\mathcal{T}$.
6. נבנה סיור TSP מ- ε ע"י קיצורי דרך.

טענה: האלגוריתם הוא 1.5-מקרב.

למה: יהי OPT המחיר של סיור TSP אופטימלי. תהי קבוצה $W \subseteq V(G)$ כך ש- $|W|$ זוגית.

ויהי M שידוך מושלם במחיר מינימום ב- $G[W]$.

אזי, $c(M) \le OPT/2$.

הוכחה: נוכיח את הטענה על האלגוריתם תחת ההנחה שהלמה נכונה. ואז נוכיח את הלמה ונסיים.

יהי L הסיור המתקבל מהאלגוריתם. אזי, העלות של L היא לכל היותר כל ה-MST, ועוד הצלעות שהוספנו בשידוך (בגלל אי"ש המשולש).

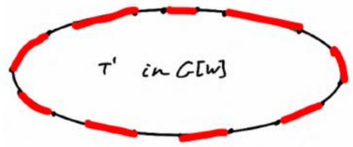
כלומר לפי הלמה ($c(M) \le OPT/2$) ולפי הטענה מההתחלה ($c(T) \le OPT$) נקבל:

$$c(L) \le c(T) + c(M) \le OPT + OPT/2 = 1.5 \cdot OPT$$

נוכיח את הלמה: יהי T סיור TSP אופטימלי, ויהי T' סיור TSP ב- $G[W]$ שהתקבל ע"י קיצורי דרך ב-T.

בגלל אי"ש המשולש, נקבל $c(T') \le c(T)$.

מכיוון ש- $|W|$ זוגי, T' הוא מעגל זוגי, כלומר הוא איחוד של שני שידוכים מושלמים (זרים בצלעות) של $G[W]$:



מכיוון ש-M הוא במחיר מינימום ב- $G[W]$, המשקל שלו הוא לכל היותר חצי המשקל של T' (כי T' מורכב משני שידוכים): $c(M) \le c(T')/2$.

אזי:

$$c(M) \le c(T')/2 \le c(T)/2 = OPT/2$$

כנדרש.