

שאלה 4: מכרזי פירסום - איכות המודעה (הקורס של טים, תרגילים 32, 34)

בשיעור הנחנו, שהסתברות ההקלקה על מודעה מסוימת תלויה רק במיקום שלה ולא במודעה. במציאות, הסתברות ההקלקה על מודעה תלויה גם במיקום וגם במודעה עצמה.

א. נניח שלכל מודעה j יש מקדם איכות q_j , והסתברות ההקלקה כששמים אותה במקום k היא:

$$r_{j,k} = r_k * q_j$$

תארו אלגוריתם המשבץ מודעות למיקומים וממקסם את סכום הערכים. הוכיחו את נכונות האלגוריתם.

ב. נניח שהסתברות ההקלקה של מודעה נתונה על-ידי פונקציה לא-ליניארית כלשהי:

$$r_{j,k} = f(r_k, q_j)$$

הראו דוגמה לפונקציה f כלשהי, שבה האלגוריתם מסעיף א לא עובד (לא מוצא שיבוץ המקסם את סכום הערכים).

פתרון - סעיף א':

טענה:

אלגוריתם חמדני ימקסם את סכום הערכים.

הוכחה:

נסדר את מקדמי האיכות של כל מודעה q_i בסדר עולה, כך שעבור n מפרסמים נקבל: $q_1 > \dots > q_n$

נמקם את המפרסם עם מקדם האיכות ה- j במיקום ה- j .

נניח בשלילה שקיים סדר אחר, לא לפי אלגוריתם חמדן, הממקסם את סכום הערכים.

זאת אומרת שקיימים לפחות שני מפרסמים, מפרסם i שנמצא במקום ה- m ומפרסם j שנמצא במקום ה- n , כך ש: $r_m < r_n$ ועבורם מתקיים: $q_i > q_j$.

נחליף את השיבוץ עבור שני המפרסמים i, j ונקבל:

$$\begin{aligned} Diff = newSum - oldSum &= [q_i * r_n + q_j * r_m] - [q_i * r_m + q_j * r_n] \\ &= q_i * (r_n - r_m) + q_j * (r_m - r_n) = (q_i - q_j) * (r_n - r_m) > 0 \blacksquare \end{aligned}$$

כיוון שהפרש סכום הערכים בסדר החדש (ע"פ אלגוריתם חמדן) ובין סכום הערכים לפי הסדר הישן גדול מ-0, קיבלנו סתירה להנחת השלילה שקיים סדר שאינו חמדן הממקסם את סכום הערכים.

פתרון - סעיף ב':

אלגוריתם חמדן (שהוצג בסעיף א') לא ימקסם את סכום הערכים אם נחשב את הערכים באמצעות

$$val_{p_i, r_m} = f(p_i, r_m) = q_i * \frac{1}{r_m}$$

אם נחליף את השיבוץ עבור שני המפרסמים i, j (כפי שעשינו בסעיף א') נקבל:

$$\begin{aligned} Diff = newSum - oldSum &= \left[q_i * \frac{1}{r_n} + q_j * \frac{1}{r_m} \right] - \left[q_i * \frac{1}{r_m} + q_j * \frac{1}{r_n} \right] \\ &= q_i * \left(\frac{1}{r_n} - \frac{1}{r_m} \right) + q_j * \left(\frac{1}{r_m} - \frac{1}{r_n} \right) = (q_i - q_j) * \left(\frac{1}{r_n} - \frac{1}{r_m} \right) < 0 \blacksquare \end{aligned}$$

האלגוריתם שהוצע בסעיף א' לא ממקסם את סכום הערכים כאשר שינינו באופן הנ"ל את הפונקציה לחישוב ערכו של מפרסם בודד, בהתאם לאיכות הערך והמיקום שלו.