

บทที่ 3

ปัญหาการคัดลอกข้อมูลที่น้อยที่สุดและความซับซ้อนของปัญหา

บทนี้จะนำเสนอปัญหาที่มีชื่อว่าปัญหาการคัดลอกข้อมูลที่น้อยที่สุด และวิเคราะห์ความยากของปัญหา ซึ่งปัญหานี้จะถูกพิสูจน์ว่าเป็นปัญหาที่ยากสำหรับคอมพิวเตอร์และไม่มีขั้นตอนวิธีสำหรับหาคำตอบของปัญหาภายในเวลาที่เหมาะสมในส่วนต่อไป

3.1 ปัญหาการคัดลอกข้อมูลที่น้อยที่สุด

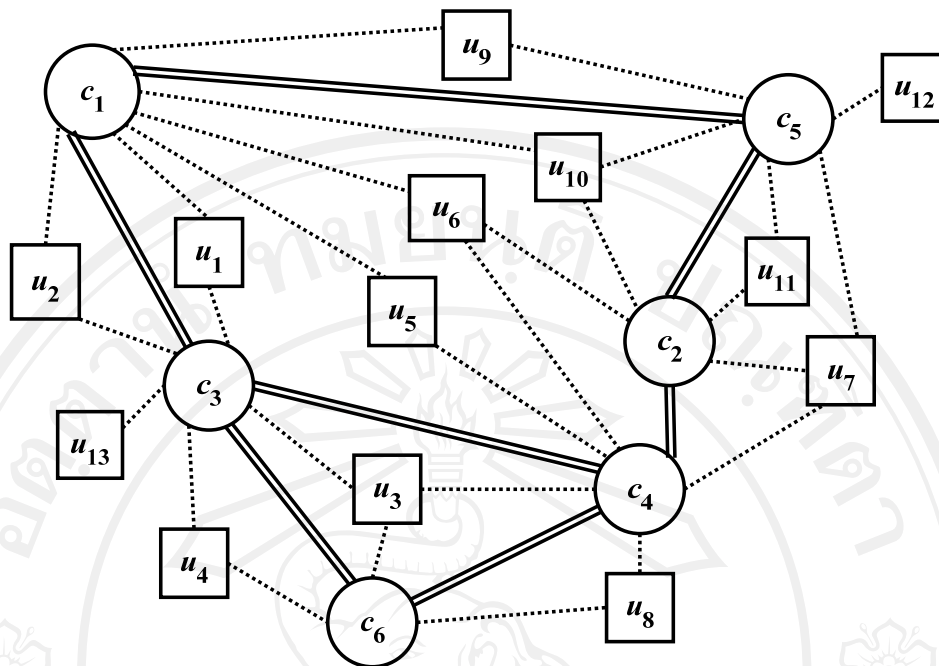
ในระบบคอมพิวเตอร์ เครื่องคอมพิวเตอร์แต่ละเครื่องจะกำหนดกลุ่มผู้ใช้ที่จะให้บริการไว้ ซึ่งกลุ่มผู้ใช้นี้จะสามารถเข้าถึงข้อมูลที่อยู่บนเครื่องคอมพิวเตอร์นั้นร่วมกันได้ ในบางครั้งอาจมีความต้องการที่จะคัดลอกข้อมูลที่อยู่บนเครื่องหนึ่งไปไว้บนเครื่องคอมพิวเตอร์เครื่องอื่น เพื่อให้ผู้ใช้ในกลุ่มสามารถเข้าถึงข้อมูลที่ต้องการ โดยที่ความต้องการจำนวนการคัดลอกข้อมูลให้น้อยครั้งที่ที่สุด หรืออีกนัยหนึ่งคือต้องการให้มีเครื่องคอมพิวเตอร์สำหรับให้บริการข้อมูลนั้นน้อยที่สุด แต่ก็ยังต้องการให้มีจำนวนผู้ใช้ที่สามารถเข้าถึงข้อมูลนั้นได้มากที่สุด โดยมีสมมติฐานว่าแต่ละเครื่องคอมพิวเตอร์มีข้อมูลที่ต้องการอยู่เพียง 1 ชุดเท่านั้น ปัญหานี้เรียกว่า *ปัญหาการคัดลอกข้อมูลที่น้อยที่สุด* (Minimum Replication Problem) หรือ เอ็มอาร์พี (MRP) ก่อนที่จะให้นิยามของปัญหากำหนดให้ฟังก์ชันการเข้าถึง (accessibility function) $\Phi: C \mapsto \mathbb{P}(U)$ เมื่อ $\mathbb{P}$ เป็นเซตกำลัง

นิยาม 3.1 ปัญหาการคัดลอกข้อมูลที่น้อยที่สุด (The Minimum Replication Problem)

กำหนด: ตัวแบบการบริหารระบบ $M = (C, L, U, P, S, O, A)$ ฟังก์ชันการเข้าถึง Φ และจำนวนเต็มบวก k และ u^*

ปัญหา: มีเซตย่อย $B \subseteq C$ โดยที่ $|B| \leq k$ และ $|\bigcup_{b \in B} \Phi(b)| \geq u^*$ หรือไม่

เพื่อให้เข้าใจนิยามของปัญหามากขึ้น พิจารณากรณีตัวอย่างอย่างง่ายสำหรับปัญหา สมมติในระบบมีคอมพิวเตอร์อยู่ 6 เครื่อง $C = \{c_1, c_2, c_3, c_4, c_5, c_6\}$ และมีผู้ใช้ 13 คน $U = \{u_1, u_2, u_3, \dots, u_{13}\}$ ดังรูปที่ 3.1 และเพื่อให้ง่ายในการทำความเข้าใจไม่สนใจสมาชิกตัวอื่นใน 7 คู่อันดับเส้นประในรูปแสดงฟังก์ชันการเข้าถึง Φ และเส้นคู่หมายถึงการเชื่อมโยงกันของคอมพิวเตอร์



รูปที่ 3.1: กรณีตัวอย่างของปัญหาการคัดลอกข้อมูลที่น้อยที่สุด

สมมติให้ $u^* = 10$ และ $k = 2$ สำหรับกรณีตัวอย่างนี้ จะทำให้คำตอบของปัญหาการตัดสินใจ (decision problem) คือ “ใช่” โดยที่เลือกเซต $B = \{c_3, c_5\}$ ซึ่งคอมพิวเตอร์ทั้งสองสามารถให้บริการกับผู้ใช้ทั้งหมด 10 คนพอดี

จากตัวอย่างอาจจะมองเห็นว่าการหาผลเฉลยของปัญหาการคัดลอกข้อมูลที่น้อยที่สุดนั้น ไม่ได้ยากมาก แต่ในกรณีที่มีคอมพิวเตอร์และผู้ใช้จำนวนมากในระบบ วิธีเดียวที่จะหาผลเฉลยของปัญหาได้ก็คือตรวจสอบทุกเซตคำตอบที่เป็นไปได้ ซึ่งจะต้องใช้เวลาในการคำนวณที่อยู่ในรูปฟังก์ชันแบบเลขชี้กำลัง (exponential running time) ของขนาดข้อมูลเข้า จะเห็นว่าปัญหาการคัดลอกข้อมูลที่น้อยที่สุดนี้เป็นปัญหาที่หาผลเฉลยได้ยากเว้นแต่คลาส $P = NP$

3.2 ความเป็นเอ็นพีคอมพลีทของปัญหาการคัดลอกข้อมูลที่น้อยที่สุด

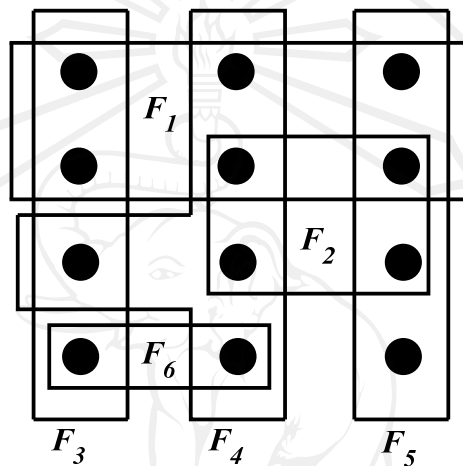
ในส่วนนี้จะพิสูจน์ว่าปัญหาการคัดลอกข้อมูลที่น้อยที่สุดเป็นปัญหาในคลาสเอ็นพีคอมพลีท (NP-complete) โดยทำการลดรูปจากปัญหา เซตคัพเวอร์ (set cover problem) หรือ เอสซีพี (SCP) ซึ่งเป็นหนึ่งใน 21 ปัญหาที่มีชื่อเสียงของ Karp ซึ่งถูกพิสูจน์ว่าเป็นปัญหาเอ็นพีคอมพลีท ในปี 1972 [11]

นิยาม 3.2 ปัญหาเซตคัฟเวอร์ (The Set Cover Problem)

กำหนด: เซตจำกัด X และเซต F ของเซตย่อยของ X โดยที่ $X = \bigcup_{Y \in F} Y$ และจำนวนเต็มบวก k^*

ปัญหา: มีเซตย่อย $D \subseteq F$ โดยที่ $|D| \leq k^*$ และ $X = \bigcup_{Y \in D} Y$ หรือไม่

กรณีตัวอย่างของปัญหาเซตคัฟเวอร์ สมมติให้ $X = \{x_1, x_2, \dots, x_{12}\}$ และเซต F กำหนดดังรูปที่ 3.2 โดยที่จุดหมายถึงสมาชิกแต่ละตัวของ X และเส้นหมายถึงขอบเขตของเซต F



รูปที่ 3.2 : กรณีตัวอย่างปัญหาเซตคัฟเวอร์

สมมติให้ $k^* = 3$ สำหรับกรณีตัวอย่างของปัญหาเซตคัฟเวอร์นี้ จะได้คำตอบคือ “ใช่” โดยที่ $D = \{F_3, F_4, F_5\}$ และเป็นเซตคำตอบที่น้อยที่สุดสำหรับกรณีตัวอย่างนี้
ลำดับต่อไปจะพิสูจน์ว่าปัญหาการคัดลอกข้อมูลที่น้อยที่สุดเป็นปัญหาในคลาสเอ็นพีคอมพลีท

ทฤษฎีบท 3.3 ปัญหาการคัดลอกข้อมูลที่น้อยที่สุดเป็นปัญหาเอ็นพีคอมพลีท พิสูจน์

จะแสดงว่าปัญหาการคัดลอกข้อมูลที่น้อยที่สุด หรือเอ็มอาร์พี (MRP) เป็นปัญหาในคลาสเอ็นพี สมมติให้เซตคำตอบ B เซตเป็นคำตอบของกรณีตัวอย่างใดๆ (instance) จะมีขั้นตอนวิธีที่ตรวจสอบว่าเซต $B \subseteq C$ โดยที่ $|B| \leq k$ และ $|\bigcup_{b \in B} \Phi(b)| \geq u^*$ ซึ่งการตรวจสอบทั้ง 3 เงื่อนไขสามารถทำได้ภายในเวลาแบบพหุนาม ดังนั้นเอ็มอาร์พีจึงอยู่ในคลาสเอ็นพี

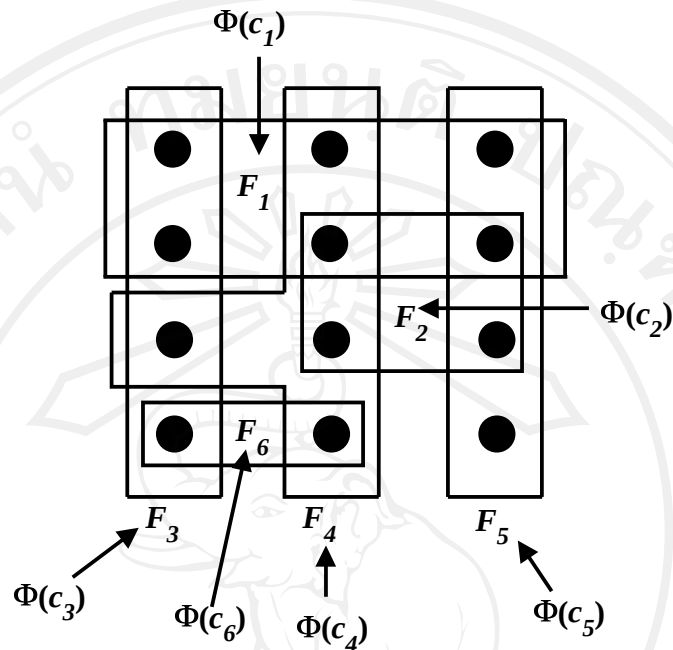
ต่อไปจะแสดงว่าปัญหาการคัดลอกข้อมูลที่น้อยที่สุดเป็นปัญหาที่มีความยากอย่างน้อยเท่ากับปัญหาเซตคัพเวอร์ หรือเอสซีพี (SCP) กำหนดกรณีตัวอย่างใดๆ ของเอสซีพี จะสร้างกรณีตัวอย่างของเอ็มอาร์พี กำหนดให้ $n = |F|$ และ $C = \{c_1, c_2, c_3, \dots, c_n\}$ สำหรับแต่ละสมาชิก $f_i \in F$ ให้ $\Phi(c_i) = f_i$ สำหรับทุก $c_i \in C$ จากนั้นให้เซต $U = X$ ต่อไปสร้างเซตของการเชื่อมโยงโดยให้ $e = n - 1$ และให้เซต $L = \{ \{c_i, c_{i+1}\} \mid 1 \leq i \leq n-1 \}$ จากนั้นกำหนดให้เซต P มีกฎซึ่งผู้ใช้แต่ละคนใน U สามารถอ่านข้อมูลในเครื่องของตนได้ ให้เซตของสถานะของระบบ $S = \{s_1, s_2, s_3, \dots, s_n\}$ โดยที่ s_i หมายถึงสายอักขระ (string) ของบิตที่แสดงสถานะของฮาร์ดดิสก์บนคอมพิวเตอร์ c_i ต่อไปให้เซต O มีเพียงตัวดำเนินการสำหรับอ่านข้อมูลเพียงตัวเดียว และ ทุก $a_i \in A$ ให้ a_i คือการกระทำของผู้ซึ่ง อ่านข้อมูลของตนเอง ณ เวลา 1 สุดท้ายกำหนดให้ $u^* = |X|$ และ $k = k^*$ การลดรูปของปัญหาดังกล่าวสามารถคำนวณได้โดยขั้นตอนวิธีที่มีเวลาในการทำงานที่เป็นฟังก์ชันพหุนาม ลำดับต่อไปจะพิสูจน์ว่า มีเซตย่อย $B \subseteq C$ โดยที่ $|B| \leq k$ และ $|\bigcup_{b \in B} \Phi(b)| \geq u^*$ ก็ต่อเมื่อมีเซตย่อย $D \subseteq F$ โดยที่ $|D| \leq k^*$ และ $X = \bigcup_{Y \in D} Y$

($\rightarrow$) กำหนดให้เซต $B \subseteq C$ โดยที่ $|B| \leq k$ และ $|\bigcup_{b \in B} \Phi(b)| \geq u^*$ เป็นเซตคำตอบของปัญหาเอ็มอาร์พี จากการลดรูปข้างต้นได้ว่า ทุกสมาชิก $b_i \in B$ มี $Y_j \in F$ ซึ่ง $\Phi(b_i) = Y_j$ เพราะฉะนั้น มีเซต $D \subseteq F$ โดยที่ $|B| = |D| \leq k = k^*$ เนื่องจากกำหนดให้ $u^* = |X|$ และ $U = X$ ดังนั้น $\bigcup_{b \in B} \Phi(b) = \bigcup_{Y \in D} Y = X$

($\leftarrow$) ให้เซต $D \subseteq F$ โดยที่ $|D| \leq k^*$ และ $X = \bigcup_{Y \in D} Y$ เป็นเซตคำตอบของปัญหาเอสซีพี จากการลดรูปข้างต้นได้ว่า แต่ละ $Y_j \in D$ จะมี $b_i \in C$ ซึ่ง $\Phi(b_i) = Y_j$ ดังนั้นจะมีเซต $B \subseteq C$ โดยที่ $|D| = |B| \leq k^* = k$ และ $\bigcup_{b \in B} \Phi(b) = \bigcup_{Y \in D} Y$ แต่เนื่องจาก $X = \bigcup_{Y \in D} Y$ ดังนั้น $|\bigcup_{b \in B} \Phi(b)| \geq |X| = u^*$

เนื่องจากปัญหาการคัดลอกข้อมูลที่น้อยที่สุดหรือเอ็มอาร์พี เป็นปัญหาที่อยู่ในคลาสเอ็นพี และปัญหานี้มีความยากอย่างน้อยเท่ากับความยากของปัญหาเซตคัพเวอร์หรือเอสซีพี ดังนั้นจึงสรุปได้ว่าปัญหาการคัดลอกข้อมูลที่น้อยที่สุดจึงเป็นปัญหาในคลาสเอ็นพีคอมพลิต $\blacksquare$

เพื่อให้เข้าใจวิธีการลดรูปจากกรณีตัวอย่างใดๆ ของปัญหาเซตคัพเวอร์ไปเป็นปัญหาการคัดลอกข้อมูลที่น้อยที่สุด พิจารณารูปที่ 3.3



รูปที่ 3.3 ตัวอย่างการลดรูปจากปัญหาเซตคัพเวอร์ไปเป็นปัญหาการคัดลอกข้อมูลที่น้อยที่สุด

แต่ละจุดก็คือผู้ใช้ $u_i \in U$ ที่ถูกกำหนดให้เท่ากับสมาชิก $x_i \in X$ ในปัญหาเซตคัพเวอร์ ดังนั้นจะเห็นว่าการเลือก $F_i \in F$ ที่ครอบคลุมสมาชิก $x_1 x_2 x_3 \dots x_6$ ในเซต X ก็คือการเลือก คอมพิวเตอร์ c_i ที่ครอบคลุมผู้ใช้ $u_1 u_2 u_3 \dots u_6$ ในเซต U นั่นเอง และในที่นี้เซต $P S O$ และ A ของตัวแบบได้กำหนดให้ขึ้นมาอย่างง่ายและสอดคล้องกับปัญหาเพื่อให้สะดวกในการศึกษาปัญหา และพิสูจน์ทฤษฎีบท

จากทฤษฎีบท 3.3 หาก $P \neq NP$ ดังนั้นจะไม่มีขั้นตอนวิธีที่มีประสิทธิภาพสำหรับหาผลเฉลยของปัญหาการคัดลอกข้อมูลที่น้อยที่สุด นั่นคือมีความเป็นไปได้อย่างมากว่าจะไม่มีขั้นตอนวิธีที่สามารถหาผลเฉลยของปัญหานี้ได้ภายในเวลาแบบพหุนาม