Project 1 Phase2:
A method to improve computing efficiency

Balanced information exposure estimation

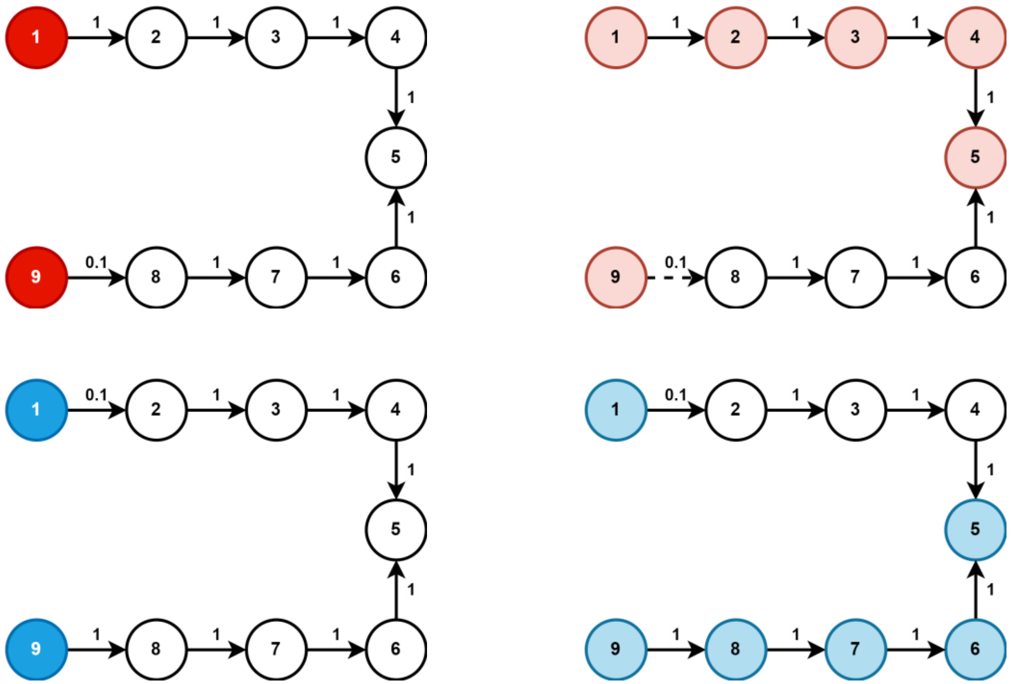
Monte Carlo simulation

- A computational algorithm that uses **repeated random sampling** to obtain the likelihood of a range of results of occurring

Estimate balanced

information exposure:

$$\Phi_{g\sim G}(S_1, S_2) = \left|V \setminus \left(r_1(I_1 \cup S_1) \triangle r_2(I_2 \cup S_2)\right)\right|_g = |\{1,2,5,8,9\}| = 5$$



Balanced information exposure estimation

Monte Carlo simulation

- A computational algorithm that uses **repeated random sampling** to obtain the likelihood of a range of results of occurring

Estimate balanced information exposure:

$$\max \Phi(S_1, S_2) = \max \mathbb{E}\big[\big|V \setminus \big(r_1(I_1 \cup S_1) \triangle r_2(I_2 \cup S_2)\big)\big|\big]$$

$$\boxed{\widehat{\Phi}(S_1, S_2) = \frac{\sum_{i=1}^{N} \Phi_{g_i}(S_1, S_2)}{N}}$$

Heuristic algorithm for IEM

Greedy best-first search

- Main idea: expand the node with the largest $h(v)$ value

$h(v)$ = increment to the balanced information exposure

Algorithm: Greedy best-first search

$S_1 \leftarrow S_2 \leftarrow \emptyset;$

while $|S_1| + |S_2| \leq k$ do

$v_1^* \leftarrow \arg\max_v \big(\Phi(S_1 \cup \{v\}, S_2) - \Phi(S_1, S_2)\big);$

$v_2^* \leftarrow \arg\max_v \big(\Phi(S_1, S_2 \cup \{v\}) - \Phi(S_1, S_2)\big);$

add the better option between $< v_1^*, \emptyset >$ and $< \emptyset, v_2^* >$ to $< S_1, S_2 >$ while respecting the budget.

Combining Monte Carlo and greed search

$S_1 \leftarrow S_2 \leftarrow \emptyset$;

while $|S_1| + |S_2| \leq k$ **do**

for j = 1 to N:

do the following Monte Carlo sampling, each sampling to calculate the h(v) value for all vertices:

1. simulate an IC model using seed set $I_1 \cup S_1$**, record the activate set** a_1 **and exposure set** r_1

2. simulate an IC model using seed set $I_2 \cup S_2$**, record the activate set** a_2 **and exposure set** r_2

3. for each v_i **in G:**

3.1 simulate an IC model base on the a_1 **and** r_1**, record the** $a_1\_v_i\_increment$ **and** $r_1\_v_i\_increment$

3.2 simulate an IC model base on the a_2 **and** r_2**, record the** $a_2\_v_i\_increment$ **and** $r_2\_v_i\_increment$

3.3 calculate and record the $h1_j(v_i)$ **=** $\Phi(S_1 \cup \{v_i\}, S_2) - \Phi(S_1, S_2)$

3.4 calculate and record the $h2_j(v_i)$ **=** $\Phi(S_1, S_2 \cup \{v_i\}) - \Phi(S_1, S_2)$

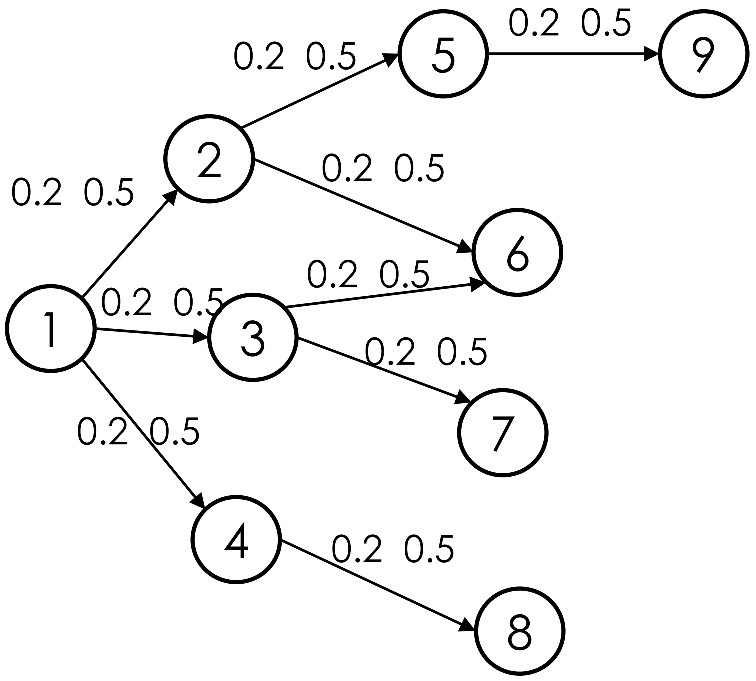
calculate the average $h1_{avg}(v)$ **value and** $h2_{avg}(v)$ **for all vertices**

$v_1^* \leftarrow \arg\max_v \left(h1_{avg}(v)\right)$;

$v_2^* \leftarrow \arg\max_v \left(h2_{avg}(v)\right)$;

add the better option between $< v_1^*, \emptyset >$ **and** $< \emptyset, v_2^* >$ **to** $< S_1, S_2 >$ **while respecting the budget.**

G:

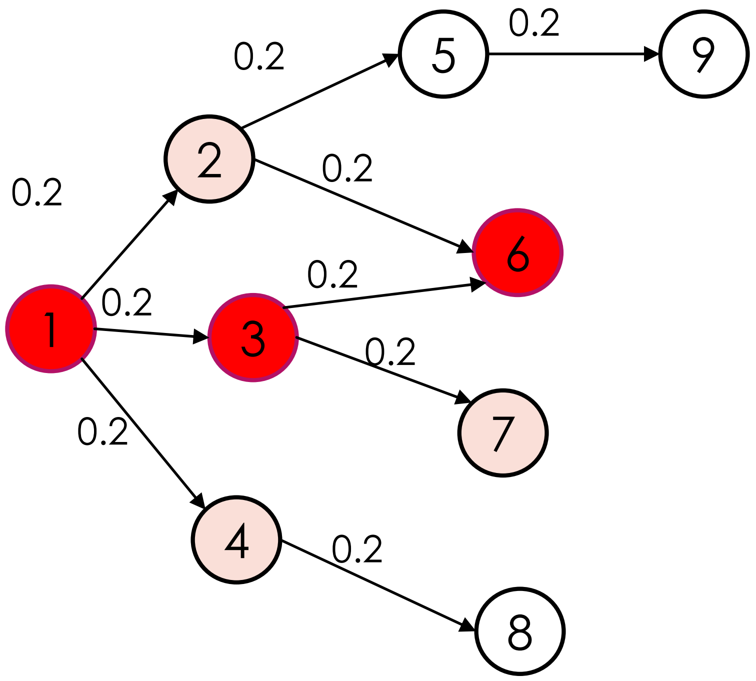


$S_1 \leftarrow S_2 \leftarrow \emptyset$

$I_1 = \{1\}$

$I_2 = \{2\}$

1. simulate an IC model using seed set $I_1 \cup S_1$, record the activate set a_1 and exposure set r_1



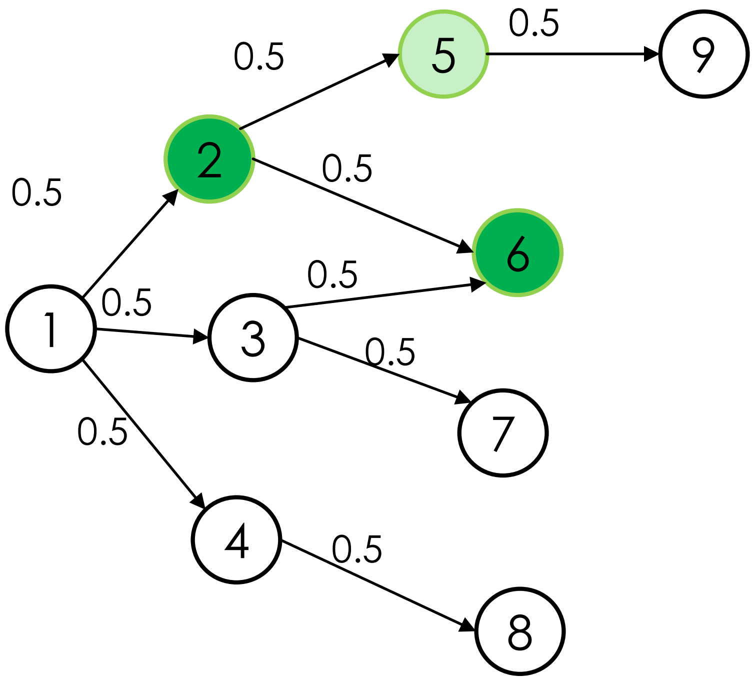
$I_1 = \{1\}$

$S_1 \leftarrow \emptyset$

	a_1	r_1
No.	IsActive	IsExposure
1	1	1
2	0	1
3	1	1
4	0	1
5	0	0
6	1	1
7	0	1
8	0	0
9	0	0

r_1={1, 2, 3, 4, 6, 7}

2. simulate an IC model using seed set $I_2 \cup S_2$, record the activate set a_2 and exposure set r_2



$I_2 = \{2\}$

$S_2 \leftarrow \emptyset$

	a_2	r_2
No.	IsActive	IsExposure
1	0	0
2	1	1
3	0	0
4	0	0
5	0	1
6	1	1
7	0	0
8	0	0
9	0	0

r_2={2, 5, 6}

3. for each v_i in G:

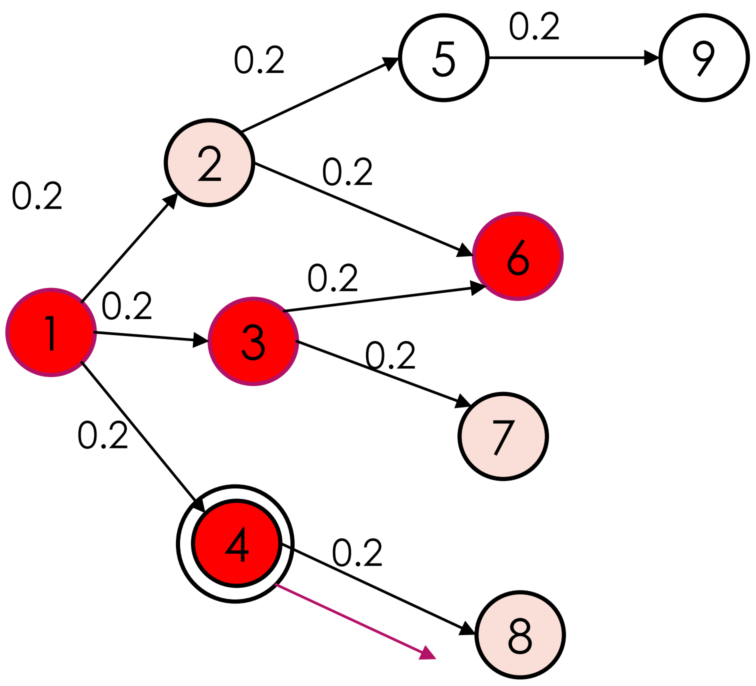
3.1 simulate an IC model base on the a_1 and r_1, record the $a_1\_v_i\_increment$ and $r_1\_v_i\_increment$

3.2 simulate an IC model base on the a_2 and r_2, record the $a_2\_v_i\_increment$ and $r_2\_v_i\_increment$

3.3 calculate and record the $h1_j(v_i) = \Phi(S_1 \cup \{v_i\}, S_2) - \Phi(S_1, S_2)$

3.4 calculate and record the $h2_j(v_i) = \Phi(S_1, S_2 \cup \{v_i\}) - \Phi(S_1, S_2)$

Take v4 for example:



	a_1	r_1
No.	IsActive	IsExposure
1	1	1
2	0	1
3	1	1
4	0	1
5	0	0
6	1	1
7	0	1
8	0	0
9	0	0

$I_1 = \{1\}$

$S_1 \leftarrow \emptyset$

$\boldsymbol{a_1\_v_4\_increment}$ **= {4}**

$\boldsymbol{r_1\_v_4\_increment}$ **= {8}**

3\. for each v_i in G:

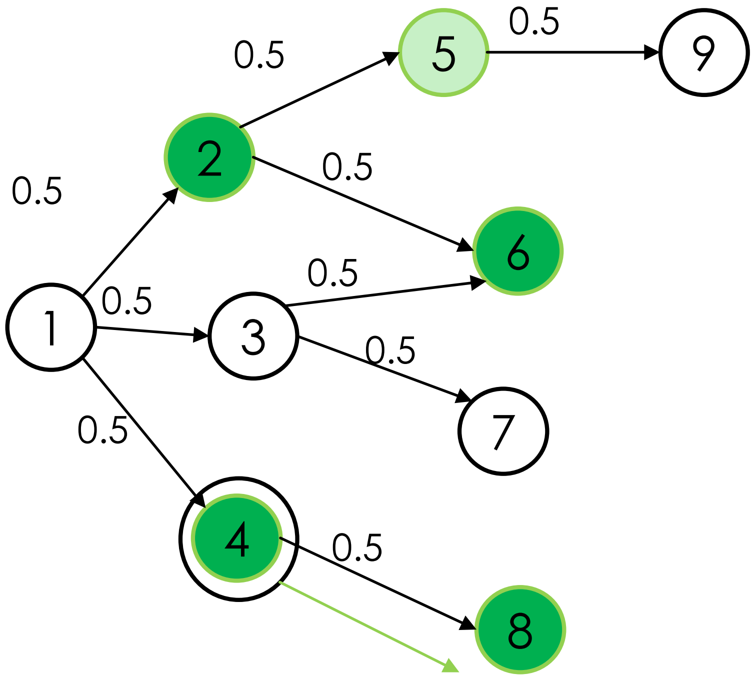
3.1 simulate an IC model base on the a_1 and r_1, record the $a_1\_v_i\_increment$ and $r_1\_v_i\_increment$

3.2 simulate an IC model base on the a_2 and r_2, record the $a_2\_v_i\_increment$ and $r_2\_v_i\_increment$

3.3 calculate and record the $h1_j(v_i) = \Phi(S_1 \cup \{v_i\}, S_2) - \Phi(S_1, S_2)$

3.4 calculate and record the $h2_j(v_i) = \Phi(S_1, S_2 \cup \{v_i\}) - \Phi(S_1, S_2)$

Take v4 for example:



	a_2	r_2
No.	IsActive	IsExposure
1	0	0
2	1	1
3	0	0
4	0	0
5	0	1
6	1	1
7	0	0
8	0	0
9	0	0

$I_2 = \{2\}$

$S_2 \leftarrow \emptyset$

$a_2\_v_4\_increment$ = {4, 8}

$r_2\_v_4\_increment$ = {4, 8}

3. for each v_i in G:

3.1 simulate an IC model base on the a_1 and r_1, record the $a_1\_v_i\_increment$ and $r_1\_v_i\_increment$

3.2 simulate an IC model base on the a_2 and r_2, record the $a_2\_v_i\_increment$ and $r_2\_v_i\_increment$

3.3 calculate and record the $\boldsymbol{h1_j(v_i)} = \boldsymbol{\Phi(S_1 \cup \{v_i\}, S_2) - \Phi(S_1, S_2)}$

3.4 calculate and record the $\boldsymbol{h2_j(v_i)} = \boldsymbol{\Phi(S_1, S_2 \cup \{v_i\}) - \Phi(S_1, S_2)}$

$(r_1=\{1, 2, 3, 4, 5, 6, 7\} \ \Delta\ r_2=\{2, 5, 6\}) = \{1, 3, 4, 7\}$

$\boldsymbol{\Phi(S_1, S_2)} = \left|V \setminus \left(r_1(I_1 \cup S_1) \triangle r_2(I_2 \cup S_2)\right)\right|_g = |V - \{1,3,4,7\}| = |\{2,5,6,8,9\}| = 5$

$\boldsymbol{r_1\_v_4\_increment} = \{8\}$

$\boldsymbol{\Phi(S_1 \cup \{v_4\}, S_2)} = \left|V \setminus \left(r_1(I_1 \cup S_1 \cup \{\boldsymbol{v_4}\}) \triangle r_2(I_2 \cup S_2)\right)\right|_g = |V \setminus (r_1 \cup \boldsymbol{r_1\_v_4\_increment} \triangle r_2)| = |V - \{1,3,4,7,8\}| = |\{2,5,6,9\}| = 4$

$\boldsymbol{h1_1(v_4)} = \boldsymbol{\Phi(S_1 \cup \{v_4\}, S_2) - \Phi(S_1, S_2) = -1}$

$\boldsymbol{r_2\_v_4\_increment} = \{4, 8\}$

$\boldsymbol{\Phi(S_1, S_2 \cup \{v_i\})} = \left|V \setminus \left(r_1(I_1 \cup S_1) \triangle r_2(I_2 \cup S_2 \cup \{\boldsymbol{v_4}\})\right)\right|_g = |V \setminus (r_1 \triangle (r_2 \cup \boldsymbol{r_2\_v_4\_increment}))| = |V - \{1,3,7,8\}| = |\{2,4,5,6,9\}| = 5$

$\boldsymbol{h2_1(v_4)} = \boldsymbol{\Phi(S_1, S_2 \cup \{v_i\}) - \Phi(S_1, S_2) = 0}$

jth Monte Carlo sampling

$$h1_{avg}(v_i) = \frac{\sum_{j=1}^{N} \mathrm{h1}_j(v_i)}{N}$$

$$h2_{avg}(v_i) = \frac{\sum_{j=1}^{N} \mathrm{h2}_j(v_i)}{N}$$