

Project1

Information Exposure Maximization

Evolutionary Optimization

- **A brief review of information exposure maximization**
- A brief review of an estimation method for balanced information exposure
- An evolutionary algorithm for information exposure maximization
- Summary

Brief review of IEM

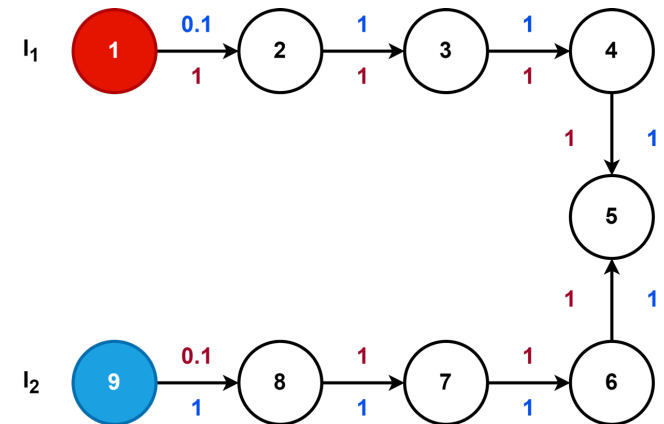
Given a social network $G = (V, E)$, two initial seed sets I_1 and I_2 , and a budget k .

The IEM is **to find two balanced seed sets S_1 and S_2** , where $|S_1| + |S_2| \leq k$, and **maximize the balanced information exposure**, i.e.,

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

$$\text{s. t. } |S_1| + |S_2| \leq k$$

$$S_1, S_2 \subseteq V$$



- Finding an optimal solution of IEM is NP-hard.
- Computing the balanced information exposure for a given solution is NP-hard.

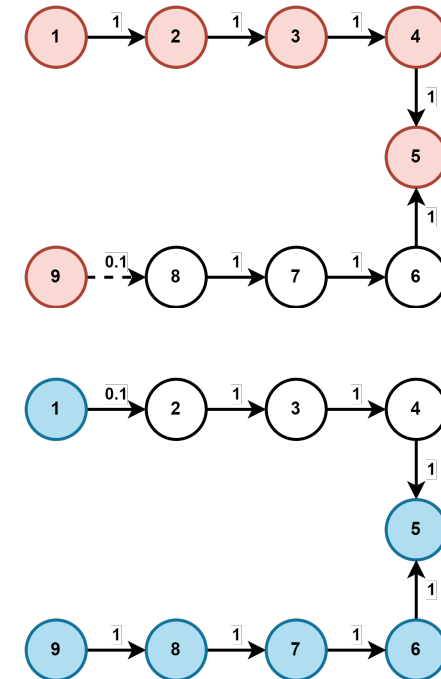
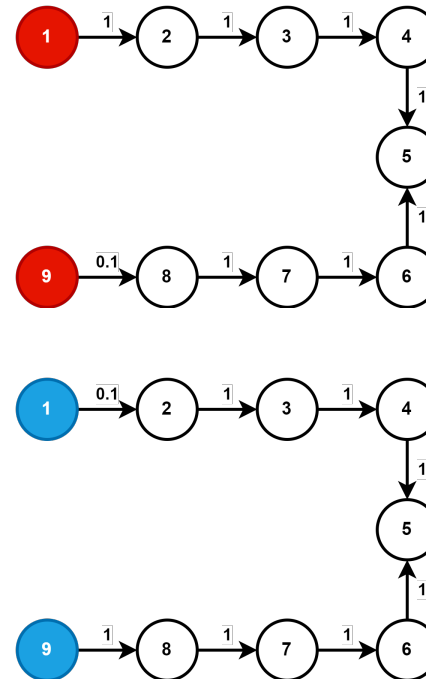
- A brief review of information exposure maximization
- **A brief review of an estimation method for balanced information exposure**
- An evolutionary algorithm for information exposure maximization
- Summary

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:

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



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:

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



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

- A brief review of information exposure maximization
- A brief review of an estimation method for balanced information exposure
- **An evolutionary algorithm for information exposure maximization**
- Summary

Basic issues

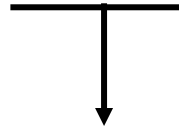
- Solution representation
 - e.g. continuous, discrete (binary, integer, permutation, etc.)
- Fitness function
 - differ from the objective function
- Search method
 - e.g., simulated annealing, evolutionary algorithms, etc.

Solution Representation

- Binary representation

$$x = \{x_1, x_2, \dots, x_{|V|}, x_{|V|+1}, x_{|V|+2}, \dots, x_{|V|+|V|}\}$$

$$x_i \in \{False, True\}$$



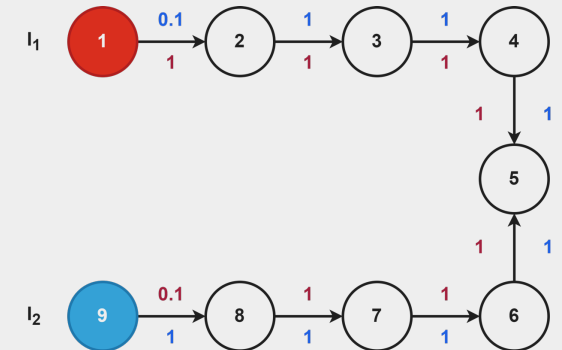
i th node is added into S_1 , $i \in [1, |V|]$

i th node is added into S_2 , $i \in [|V + 1|, |V| + |V|]$

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

$$\text{s. t. } |S_1| + |S_2| \leq k$$

$$S_1, S_2 \subseteq V$$



Fitness Function

- Distinguish between feasible and infeasible solutions

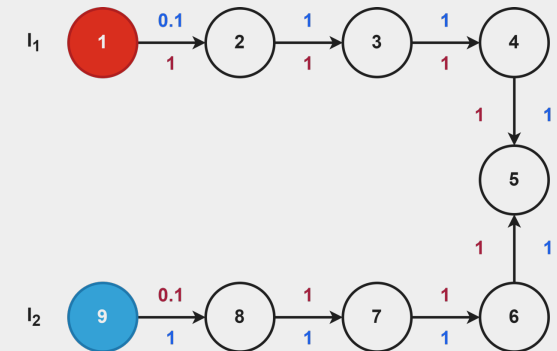
$$fitness(S_1, S_2) = \begin{cases} \hat{\Phi}(S_1, S_2) & \text{if } |S_1| + |S_2| \leq k, \\ -(|S_1| + |S_2|) & \text{otherwise.} \end{cases}$$

Punish infeasible solutions
according to the degree of violation

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

$$\text{s. t. } |S_1| + |S_2| \leq k$$

$$S_1, S_2 \subseteq V$$



Evolutionary optimization for IEM

Fitness Function (differs from objective function)

- Distinguish between feasible and infeasible solutions

$$fitness(x) = \begin{cases} \hat{\Phi}(x) & \text{if } \sum x \leq k, \\ -\sum x & \text{otherwise.} \end{cases}$$

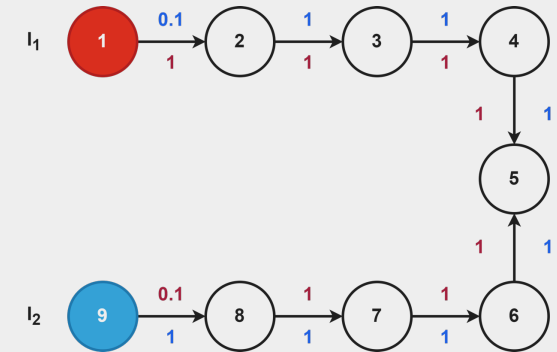
$$x = \{x_1, x_2, \dots, x_{|V|}, x_{|V|+1}, x_{|V|+2}, \dots, x_{|V|+|V|}\}$$

$$x_i \in \{False, True\}$$

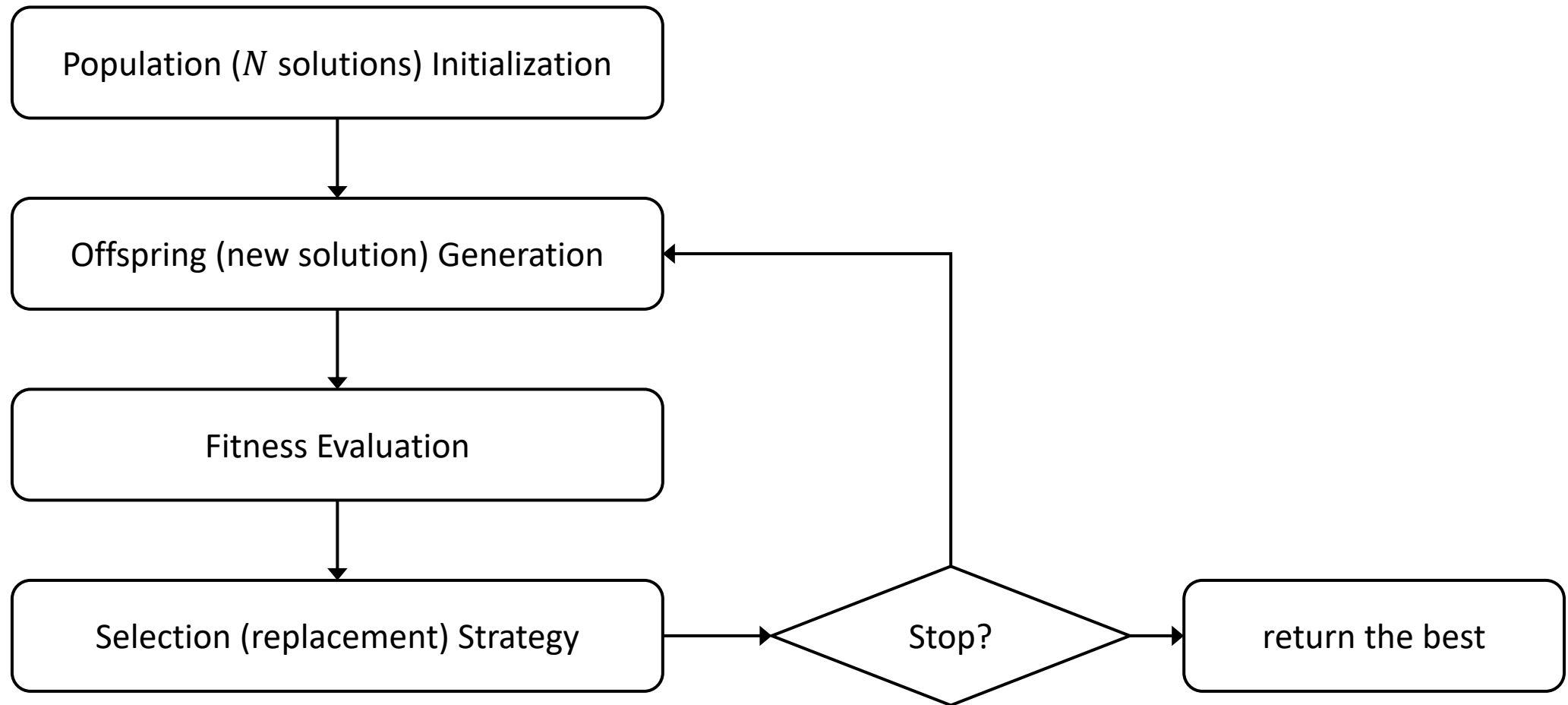
$$\max \Phi(S_1, S_2) = \max \mathbb{E}[|V \setminus (r_1(I_1 \cup S_1) \Delta r_2(I_2 \cup S_2))|]$$

$$\text{s. t. } |S_1| + |S_2| \leq k$$

$$S_1, S_2 \subseteq V$$

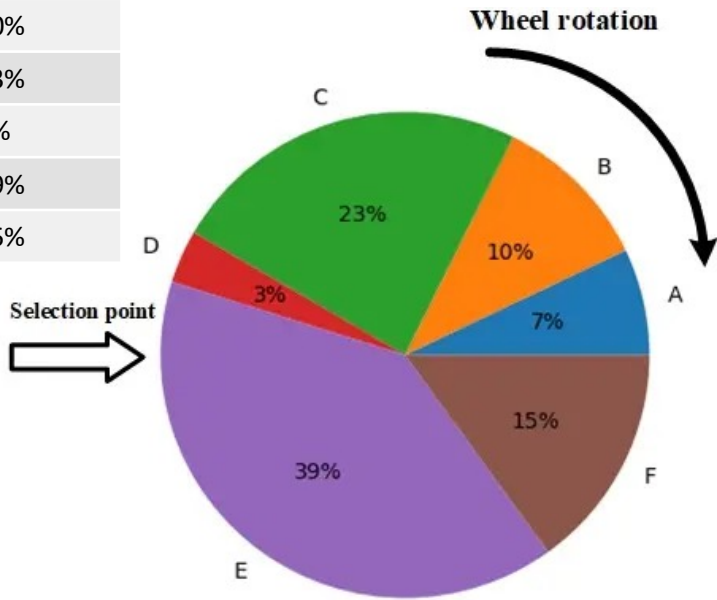


General Evolutionary Optimization Framework



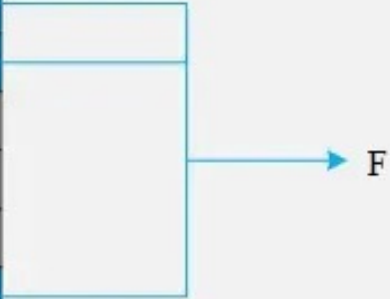
Parent selection

Solution	Fitness	proportion
A	8	7%
B	12	10%
C	27	23%
D	4	3%
E	45	39%
F	17	15%



roulette wheel selection

Individual	Fitness
A	8
B	12
C	27
D	4
E	45
F	18

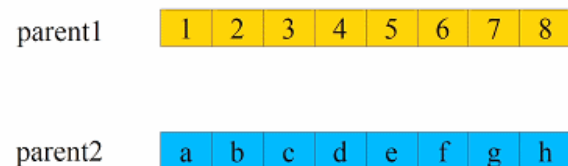


知乎 @盼小辉

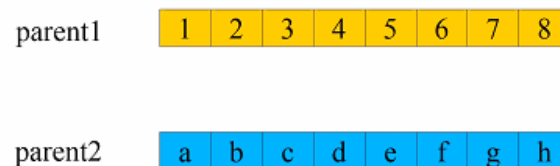
Tournament selection

Crossover

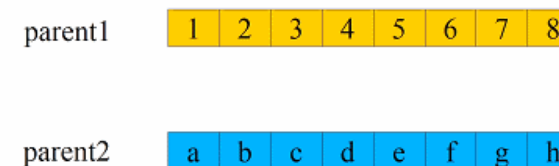
单点交叉 (Single-point crossover)



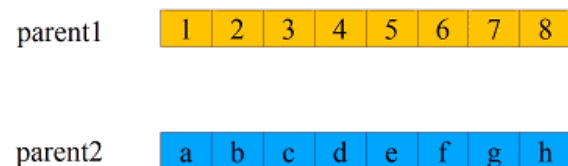
两点交叉 (Two-points crossover)



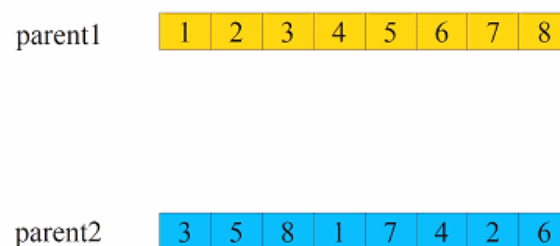
多点交叉 (Multi-point crossover)



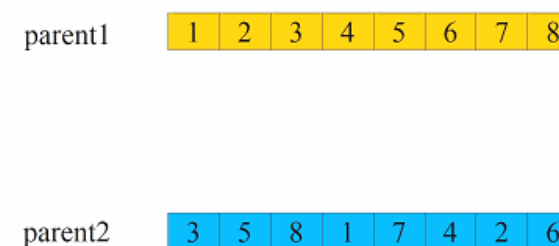
均匀交叉 (Uniform crossover)



顺序交叉 (Order crossover, OX)



位置交叉 (Position-based crossover, PBX)



Evolutionary optimization for IEM

Mutation



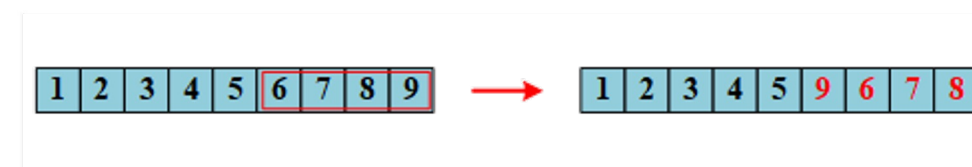
Flip bit mutation



Swap mutation



Exchange mutation



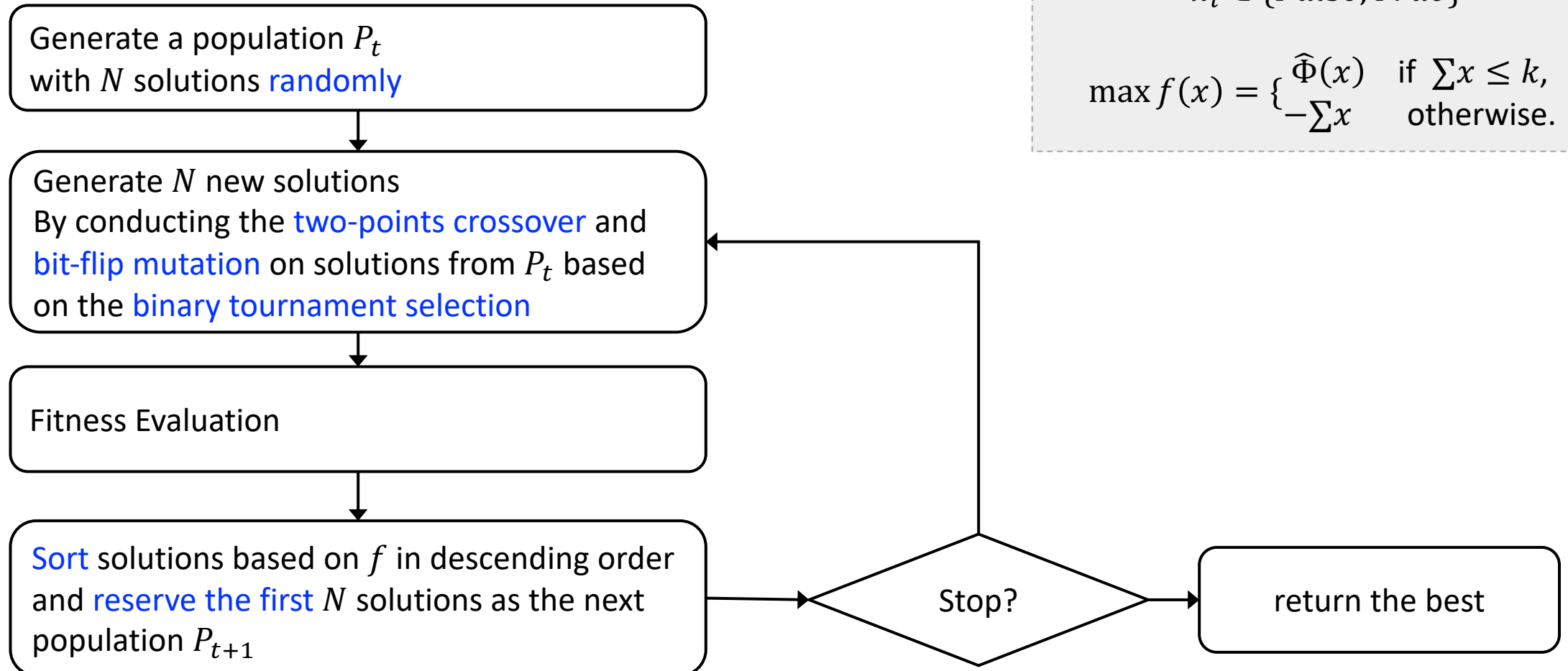
Inversion mutation

Selection strategy

Elite selection	The larger the fitness, the better
-----------------	------------------------------------

Non-Elite selection	Not entirely dependent on the fitness
---------------------	---------------------------------------

An Evolutionary Algorithm for IEM



$$x = \{x_1, \dots, x_{|V|}, x_{|V|+1}, \dots, x_{|V|+|V|}\}$$

$$x_i \in \{False, True\}$$

$$\max f(x) = \begin{cases} \hat{\Phi}(x) & \text{if } \sum x \leq k, \\ -\sum x & \text{otherwise.} \end{cases}$$

- A brief review of information exposure maximization
- A brief review of an estimation method for balanced information exposure
- An evolutionary algorithm for information exposure maximization
- **Summary**

- Information exposure maximization is computationally complex
- **Monte Carlo simulations** for balanced information exposure estimation
- **Evolutionary optimization** to find balanced seed sets
- **Improvements in solution quality or computing efficiency are encouraged**