

# Mathematical formulation

Constraints to be added to the previous formulation:

- each cluster cannot exceed a certain cardinality:

$$\forall k \leq K \quad \sum_{u \in V} x_{uk} \leq C$$

- vertices having different color cannot be clustered together:

$$\forall u \neq v \in V, k \neq l \leq K, \quad x_{uk}x_{vl} \leq \gamma_{uv}$$

where  $\gamma_{uv}$  is a new parameter defined as:

$$\forall u, v \in V, u \neq v, \quad \gamma_{uv} = \begin{cases} 1 & \text{if } u \text{ and } v \text{ have different color} \\ 0 & \text{otherwise} \end{cases}$$

# Some observations

We need to impose a maximum cardinality  $C$  for the constraint:

$$\forall k \leq K \quad \sum_{u \in V} x_{uk} \leq C$$

One possible choice is:

$$C = \lceil \frac{|V|}{2} \rceil.$$

Note that, in AMPL, we can write the constraint as:

```
subject to cardinality {k in K} :  
sum{v in V} x[v,k] <= ceil(card{V}/2);
```

# Some observations

We can allow the optimization process to determine the number of clusters actually used. We can introduce the binary variable:

$$\forall k \leq K \quad z_k = \begin{cases} 1 & \text{if cluster } k \text{ is not empty} \\ 0 & \text{otherwise} \end{cases}$$

We can change the second constraint to

$$\forall k \leq K \quad \sum_{u \in V} x_{uk} \geq z_k$$

to ensure that a cluster that does not exist need not have any vertices assigned to it, and add the term

$$\sum_{k \leq K} z_k$$

to the objective function, thus ensuring that the maximum number of clusters should be empty.