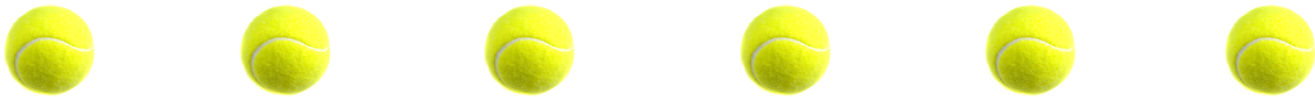


# Evaluating Adversarial Partitions

**Andreas Pashalidis & Stefan Schiffner**  
**K.U.Leuven COSIC**



ESORICS, Sep 22<sup>nd</sup> 2010, Athens

# Outline

**Motivation**

The Idea

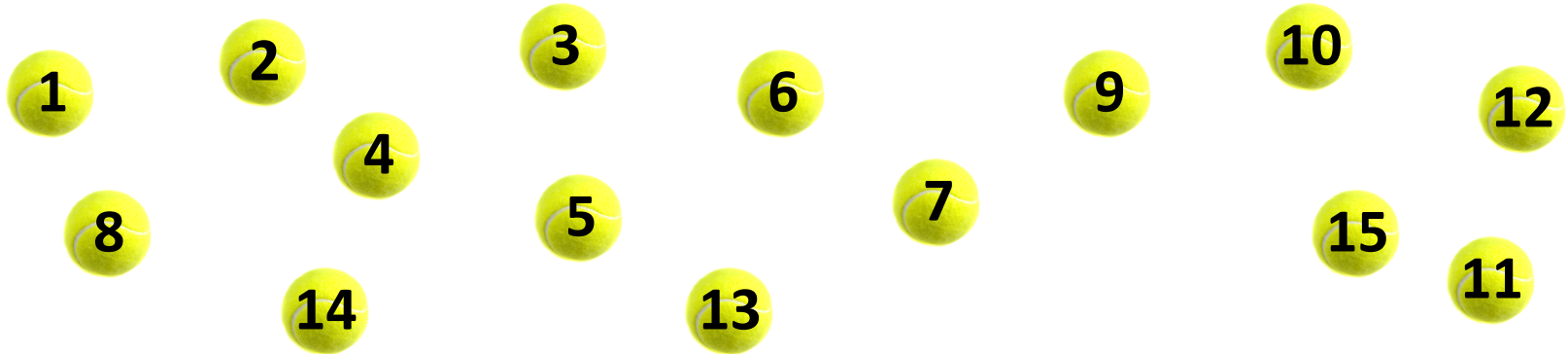
Evaluation

Conclusions

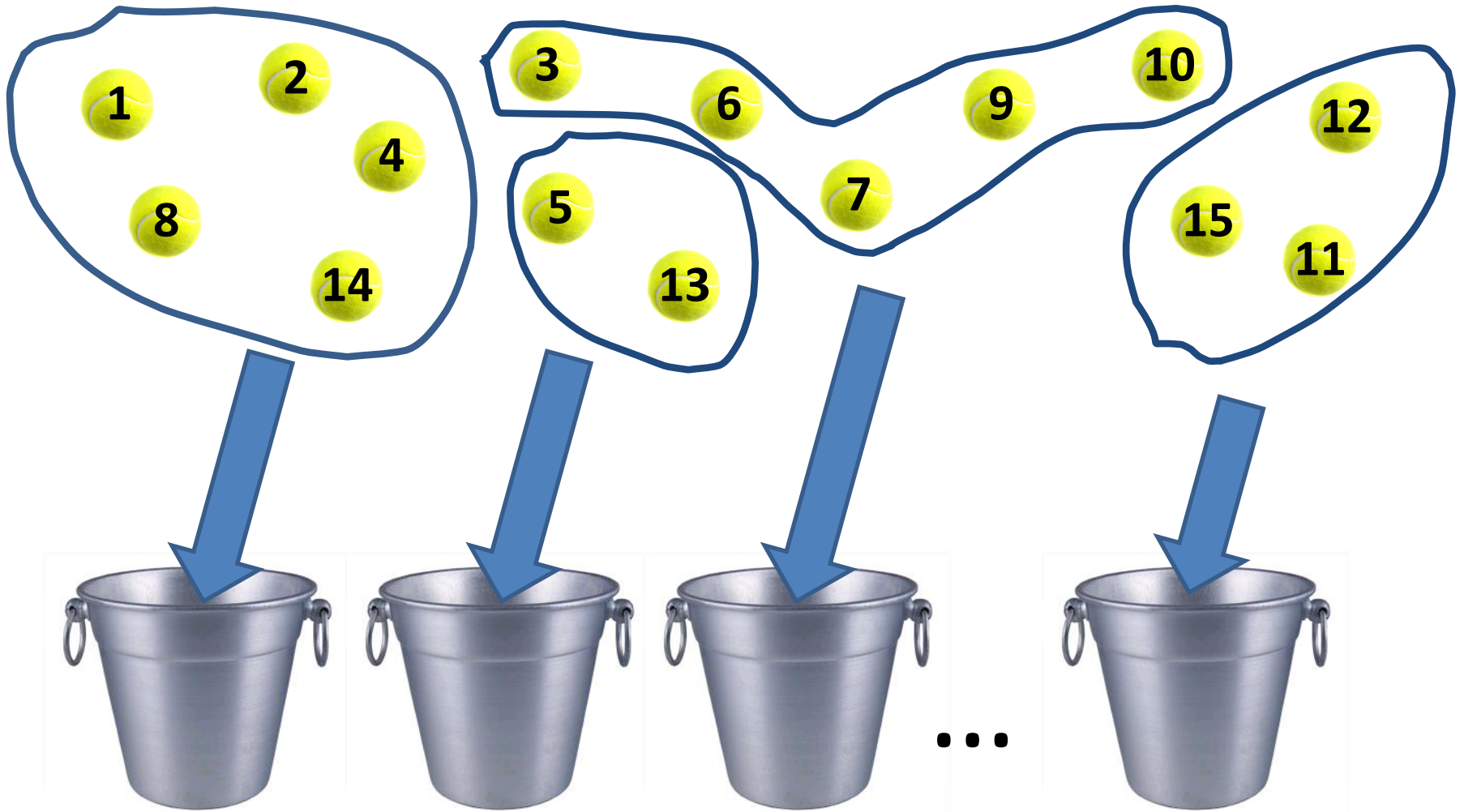
# The problem

- We would like to measure **unlinkability** of items in a privacy setting.
- Items are unlinkable when an adversary cannot say which of them belong together.
- The adversary knows the items.

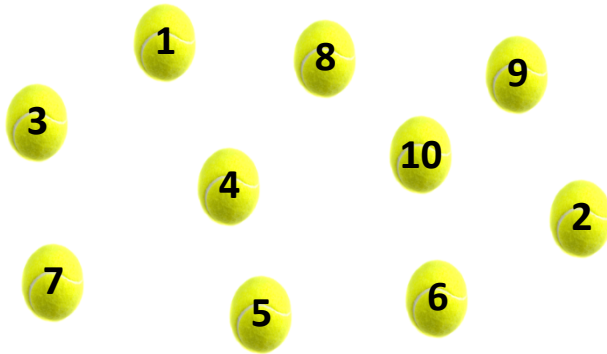
# Adversary's task



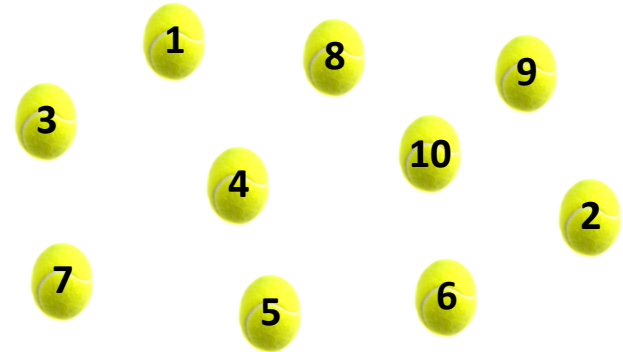
# Adversary's task



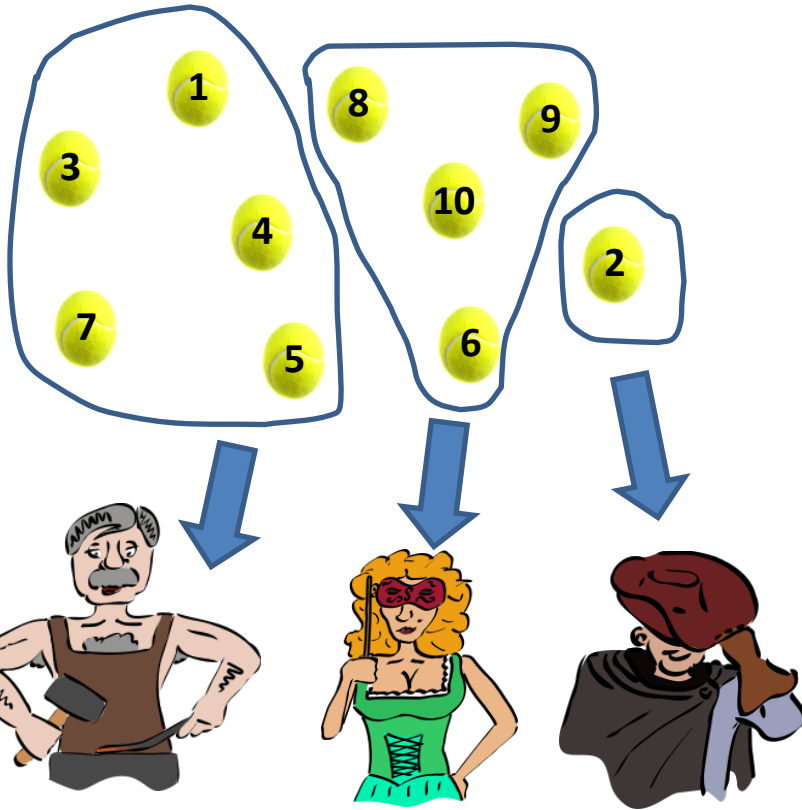
# Real World (truth)



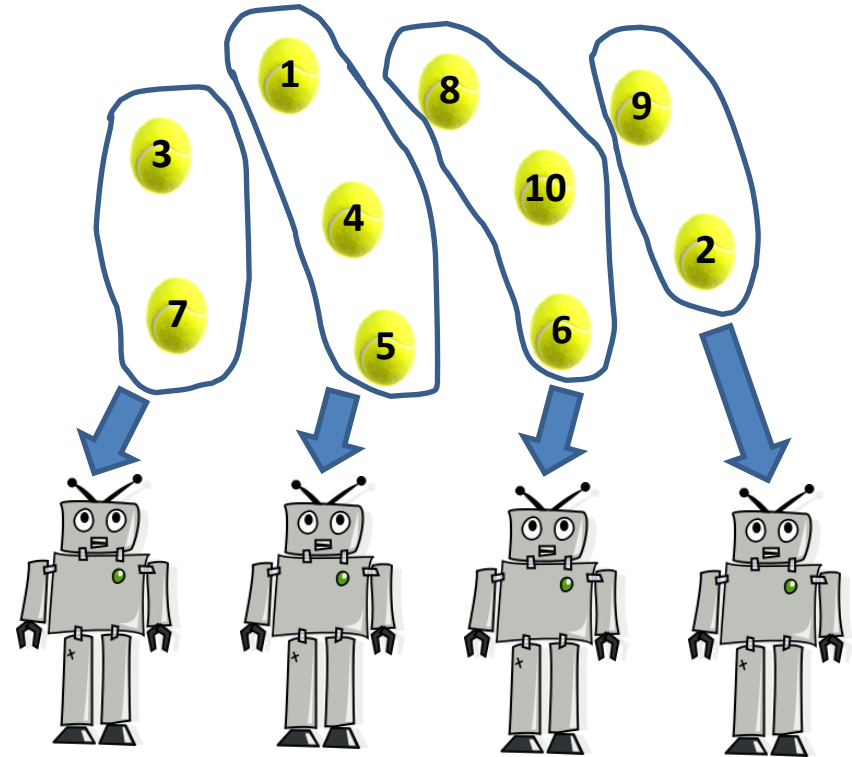
# Adversary's guess



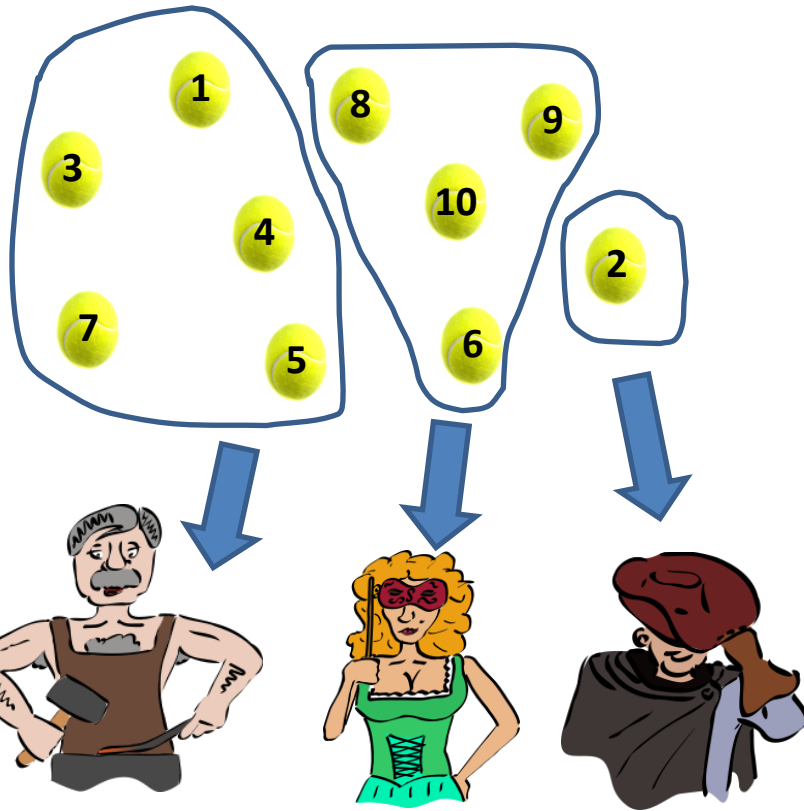
# Real World (truth)



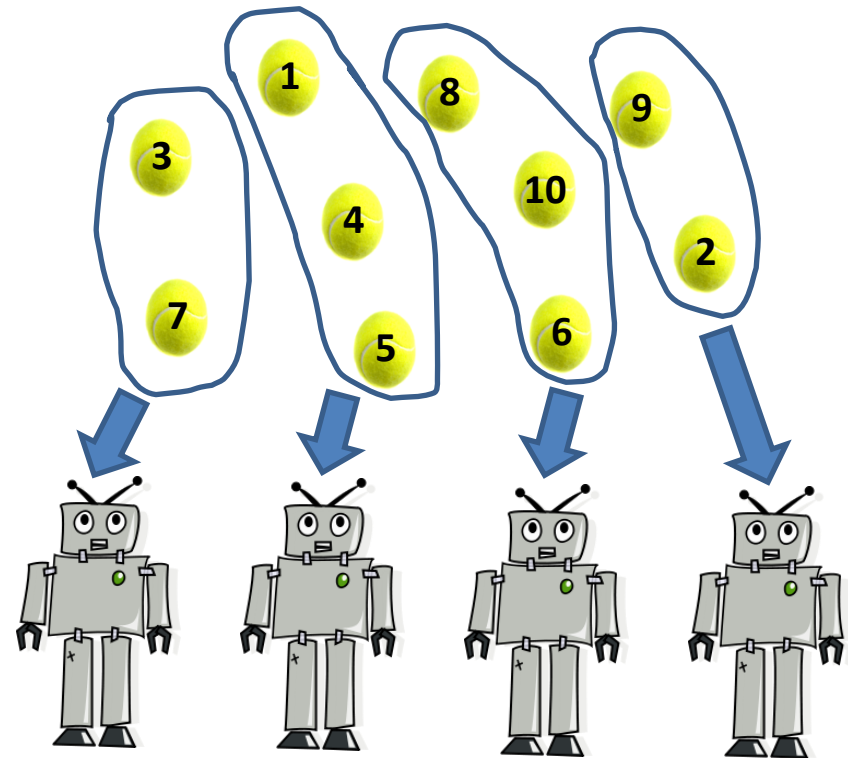
# Adversary's guess



# Real World (truth)



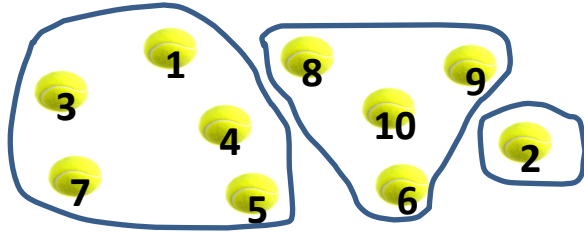
# Adversary's guess



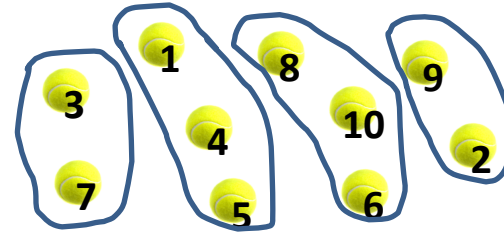
How well did the adversary do?

# Some distance measures

$$P = \{P_1, P_2, P_3\}$$

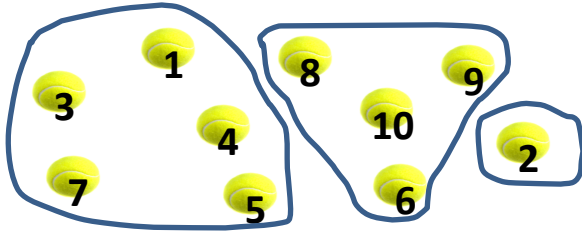


$$P' = \{P'_1, P'_2, P'_3, P'_4\}$$

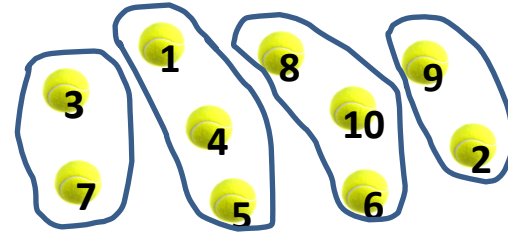


# Some distance measures

$$P = \{P_1, P_2, P_3\}$$



$$P' = \{P'_1, P'_2, P'_3, P'_4\}$$



## Rand Index

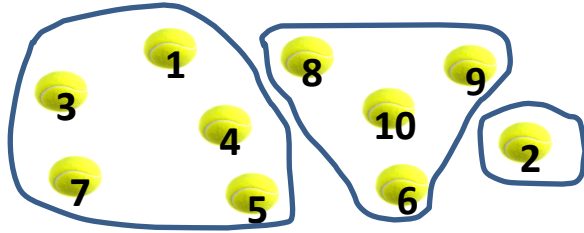


⋮

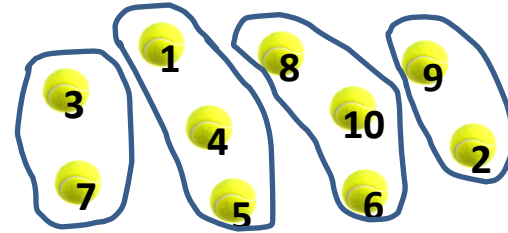


# Some distance measures

$$P = \{P_1, P_2, P_3\}$$



$$P' = \{P'_1, P'_2, P'_3, P'_4\}$$



Rand Index



⋮



**Minimum Transfer Distance**

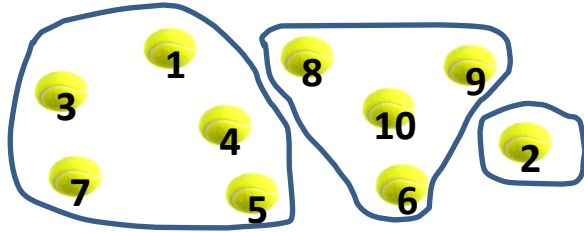
Minimum number of element reallocations to achieve equality of the two partitions.

These balls must jump;  
hence **MTD = 3**

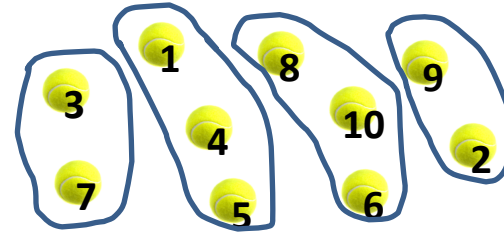


# Some distance measures

$$P = \{P_1, P_2, P_3\}$$



$$P' = \{P'_1, P'_2, P'_3, P'_4\}$$



Rand Index



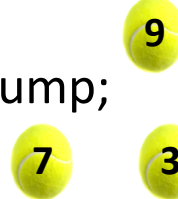
⋮



**Minimum Transfer Distance**

Minimum number of element reallocations to achieve equality of the two partitions.

These balls must jump; hence **MTD = 3**



**Variation of Information**

$$VI(P, P') = H(P) + H(P') - 2I(P, P')$$

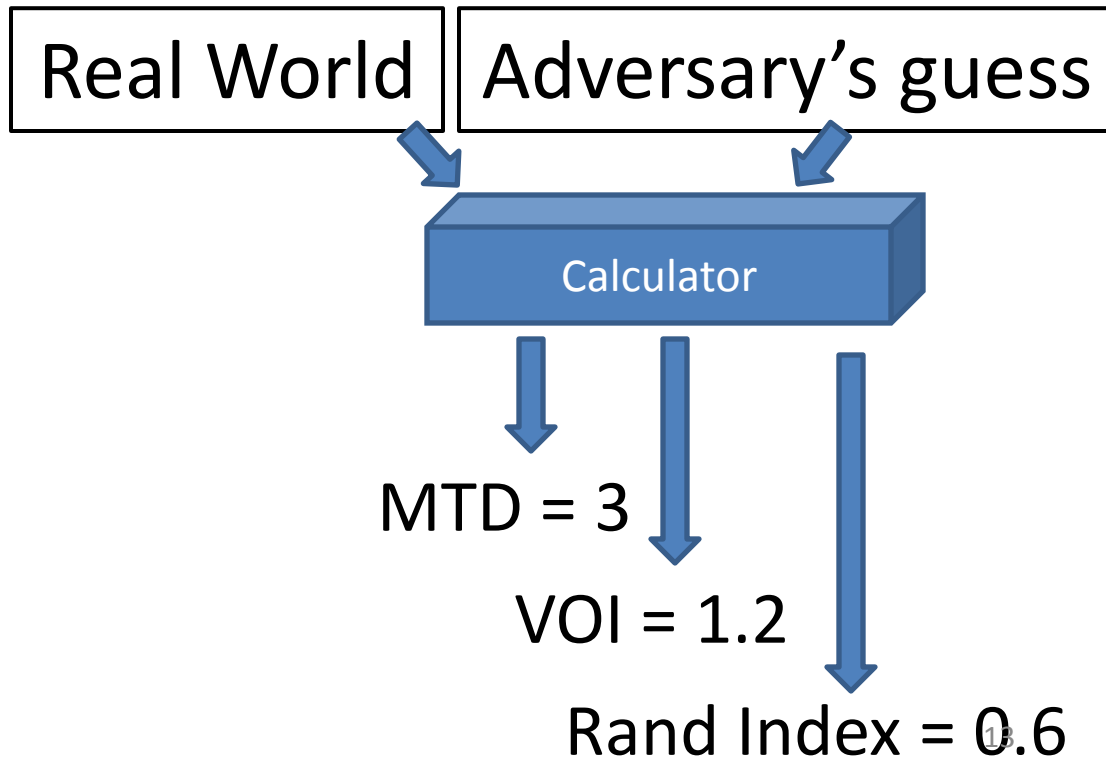
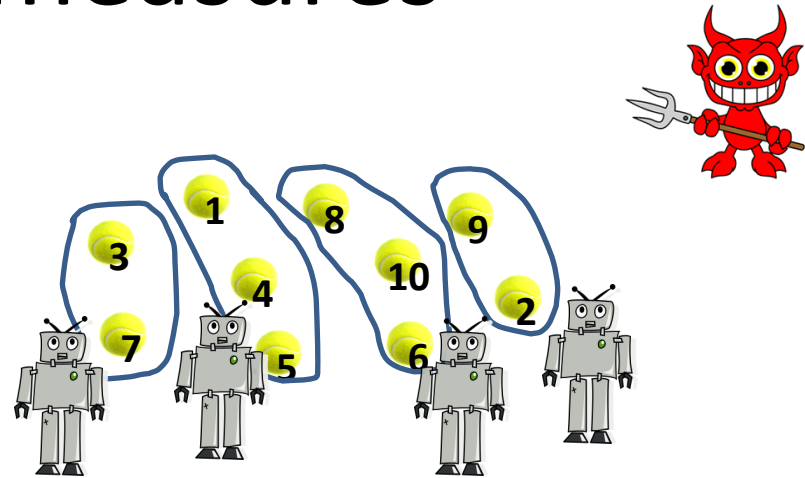
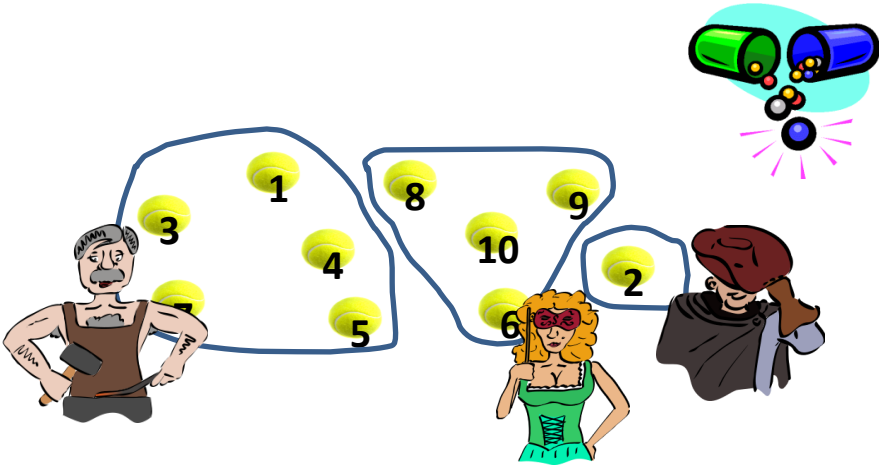
where

$$I(P, P') = \sum_{k=1}^K \sum_{k'=1}^{K'} \frac{|P_k \cap P'_{k'}|}{n} \log \frac{|P_k \cap P'_{k'}|}{nP(k)P'(k')}$$

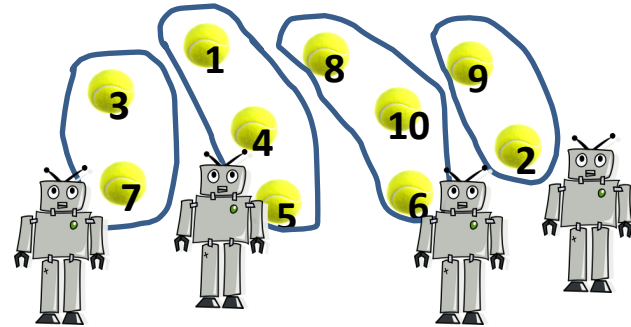
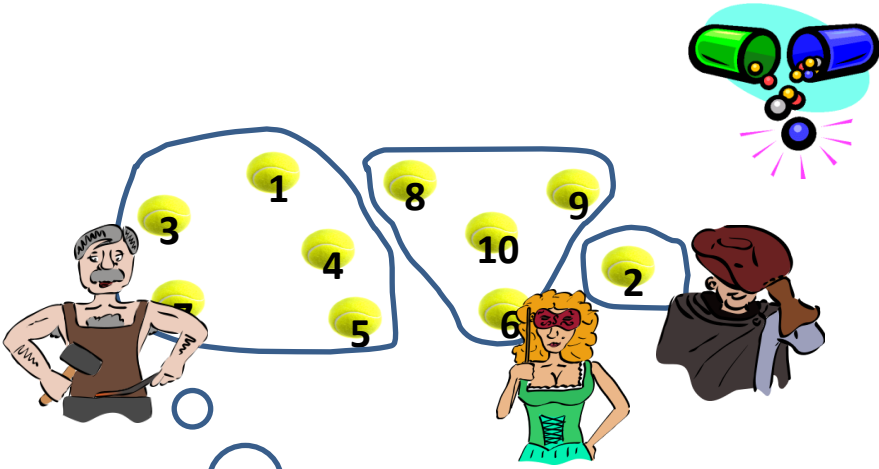
$$H(P) = -\sum_{k=1}^K P(k) \log P(k)$$

$$P(k) = |P_k| / n \quad 12$$

# Some distance measures



# Some distance measures



Are my **personal privacy preferences** taken into account?  
What does this number **mean?**

Real World

Adversary's guess

Calculator

MTD = 3

VOI = 1.2

Rand Index = 0.6

## Rephrasing the question

“How well is ***Alice’s cluster*** of items hidden within the adversarial partition?”

## Rephrasing the question

“How well is *Alice’s cluster* of items hidden within the adversarial partition?”

Alice has **preferences** and **sensitivities**.

- Privacy threshold (% of items linked)
- Linking vs. contamination
- Relative sensitivity of items

# Outline

Motivation

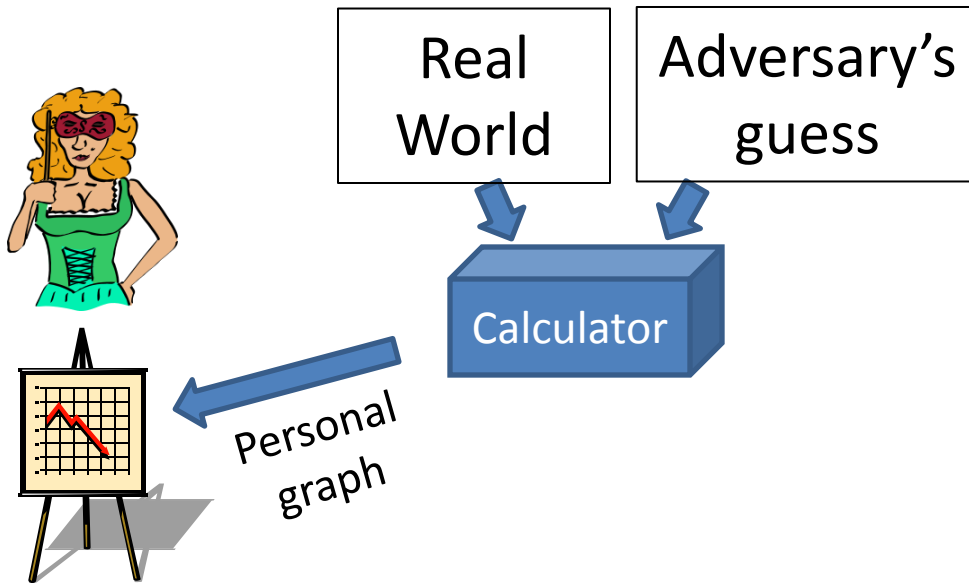
**The Idea**

Evaluation

Conclusions

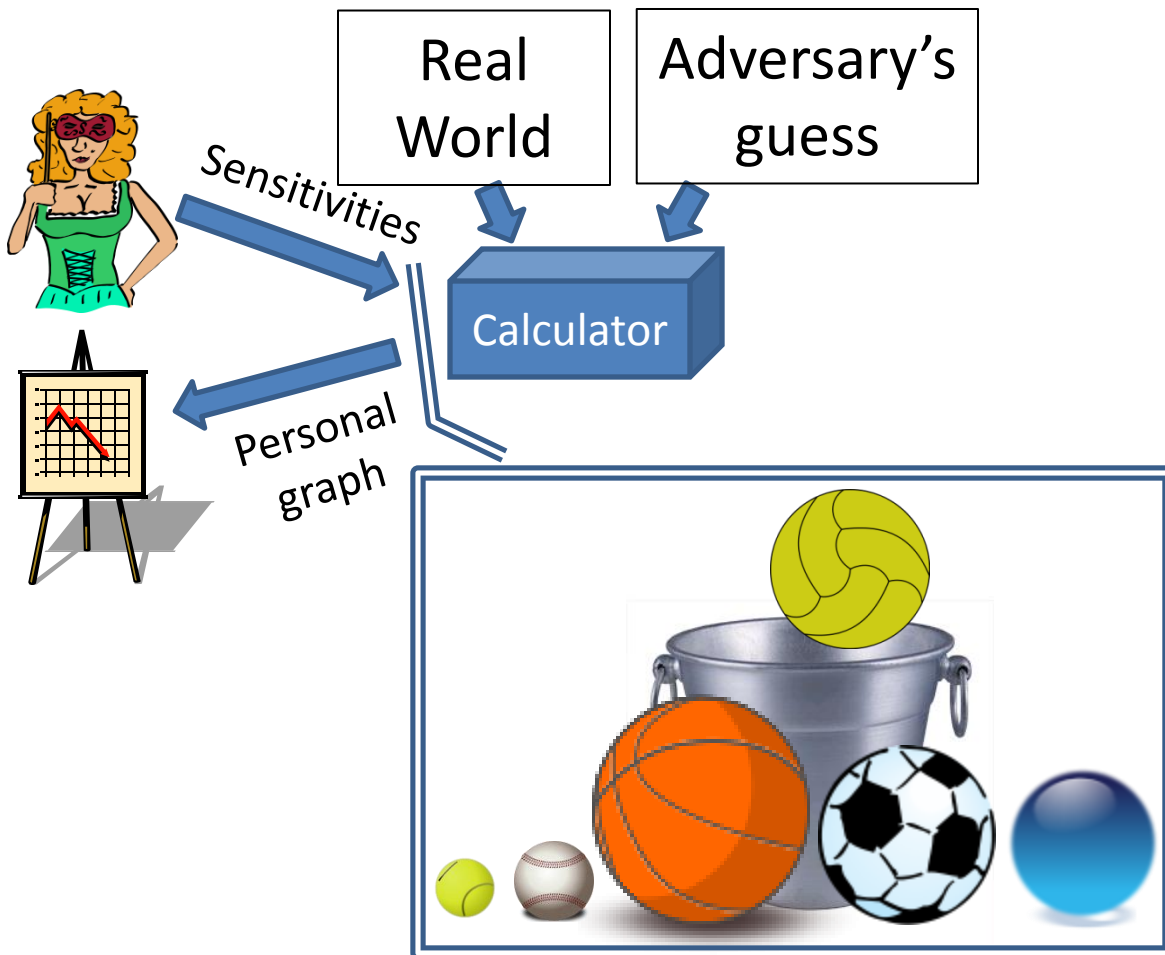
# Our idea

- For each user, draw a picture that depicts the user's **personal** situation.



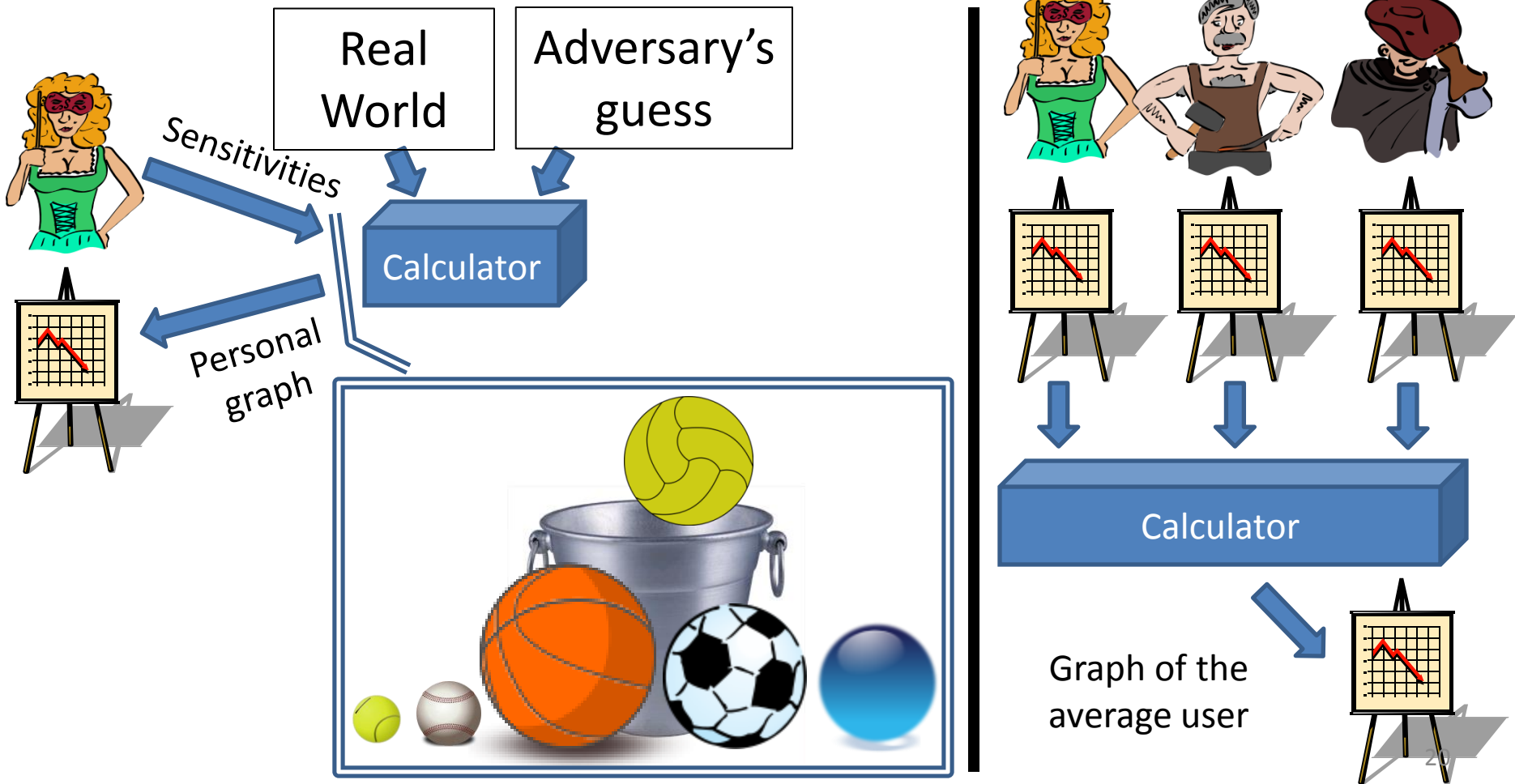
# Our idea

- For each user, draw a picture that depicts the user's **personal** situation.

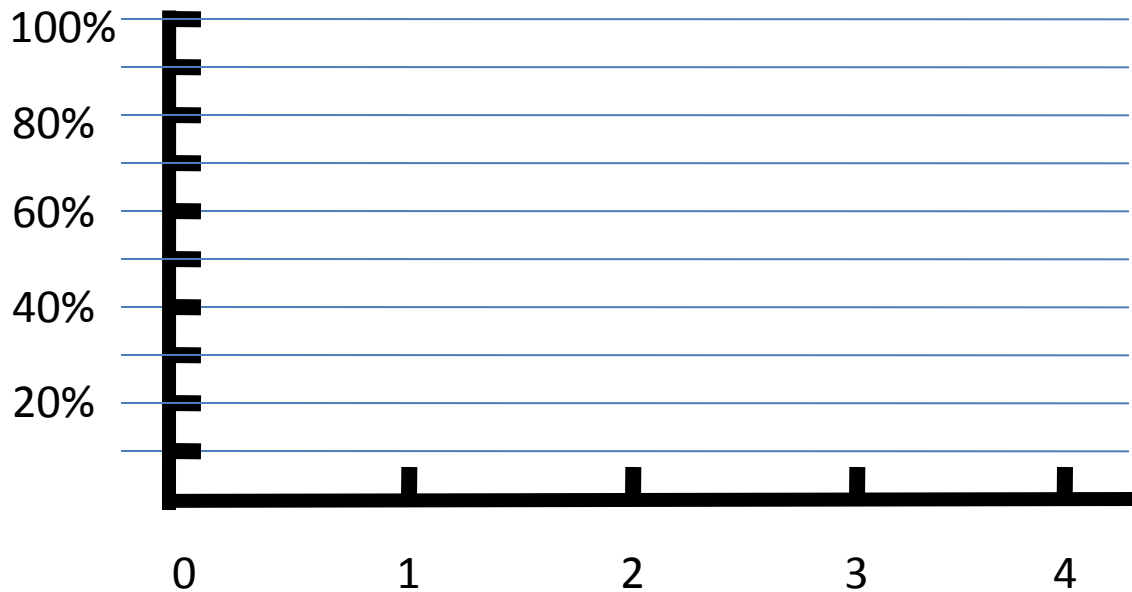
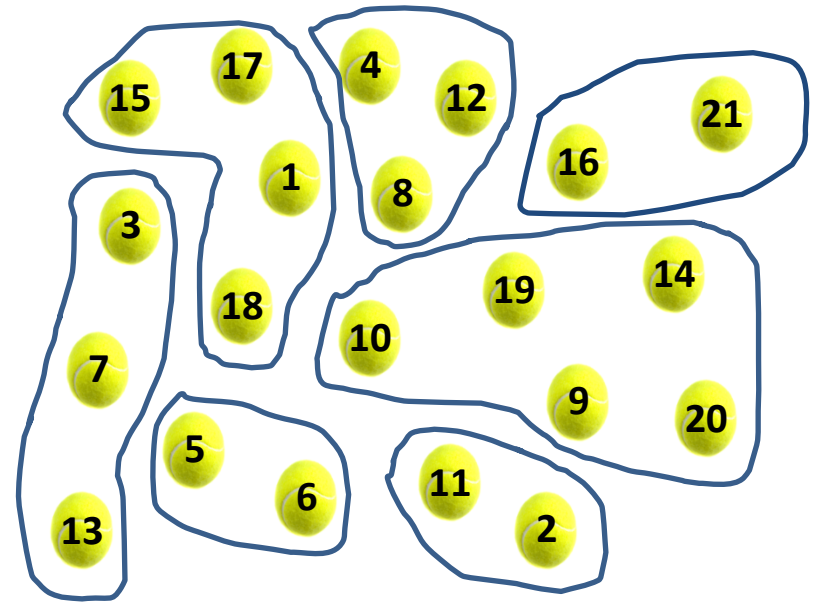
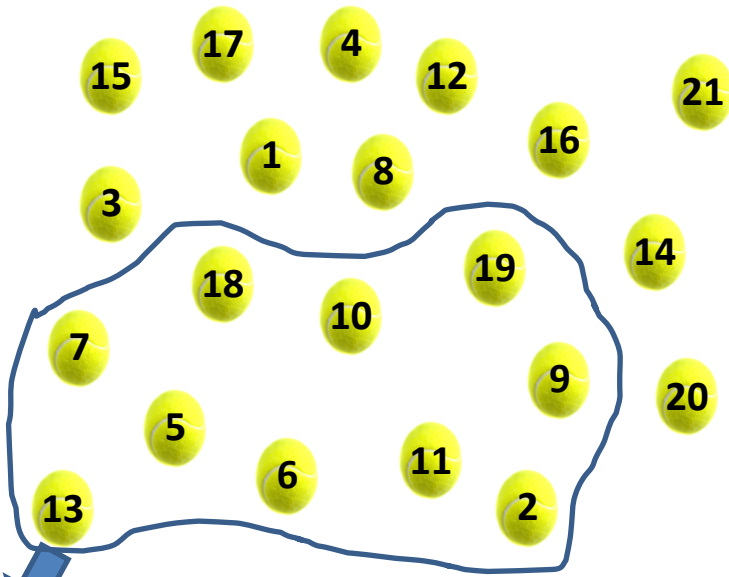


# Our idea

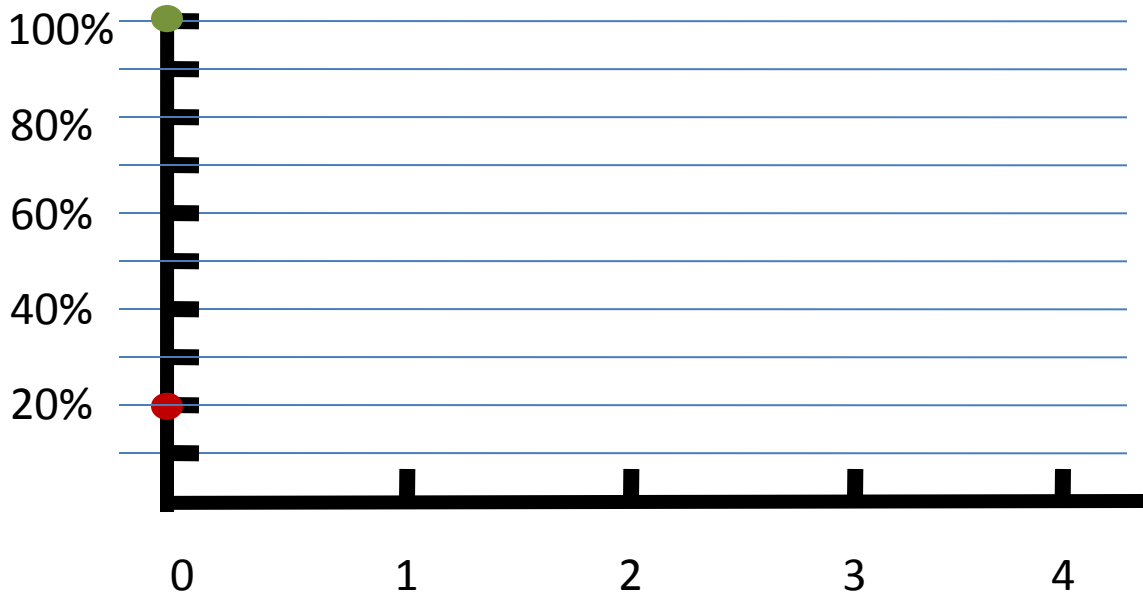
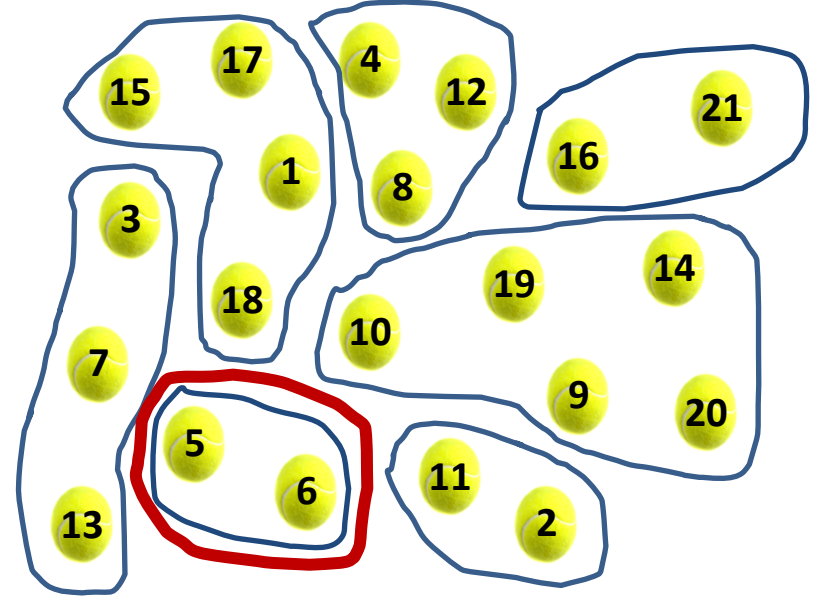
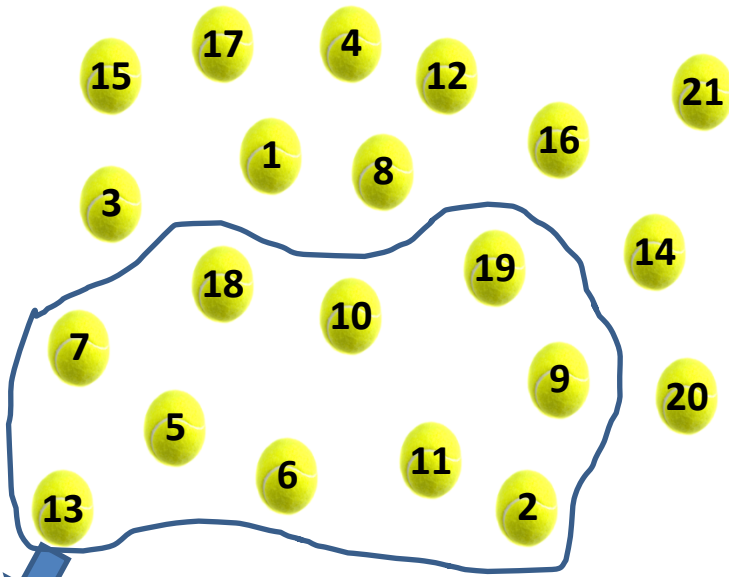
- For each user, draw a picture that depicts the user's **personal** situation.



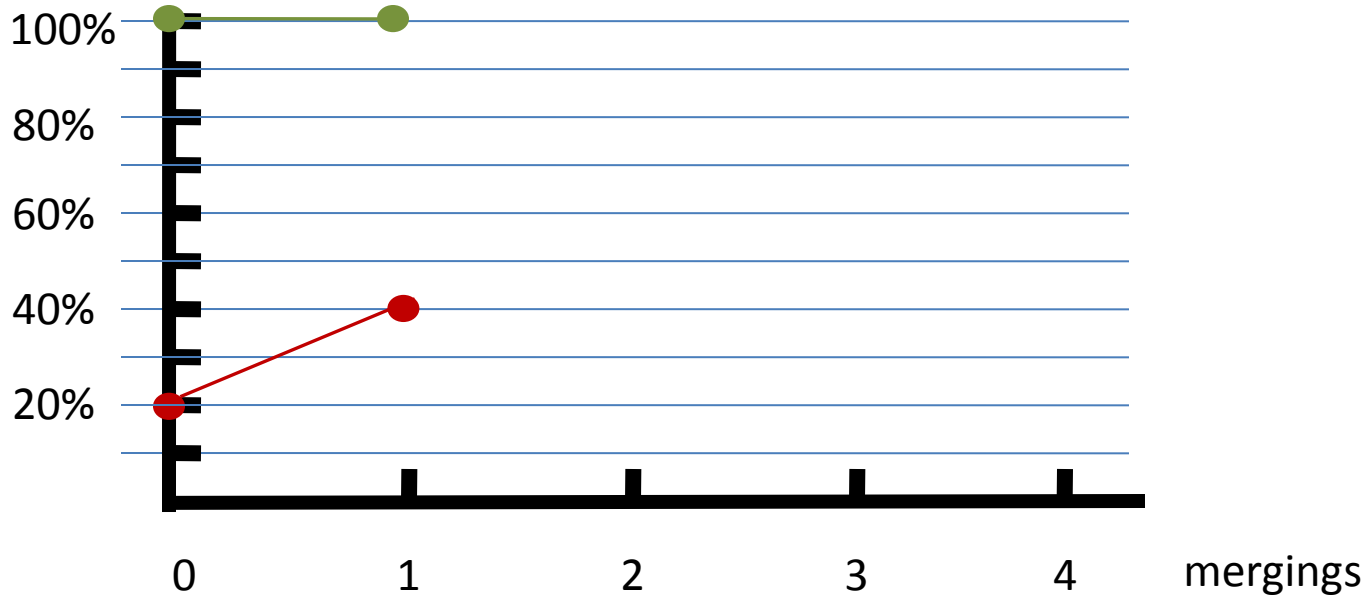
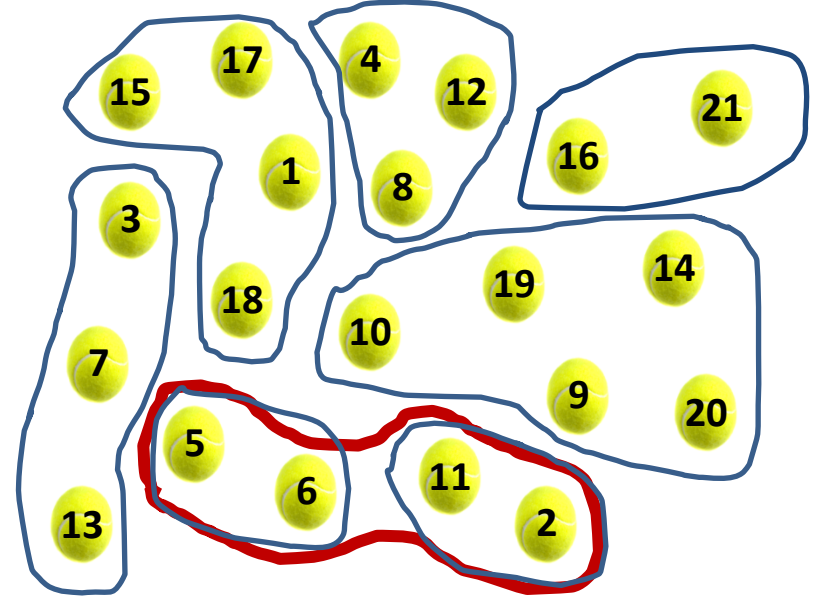
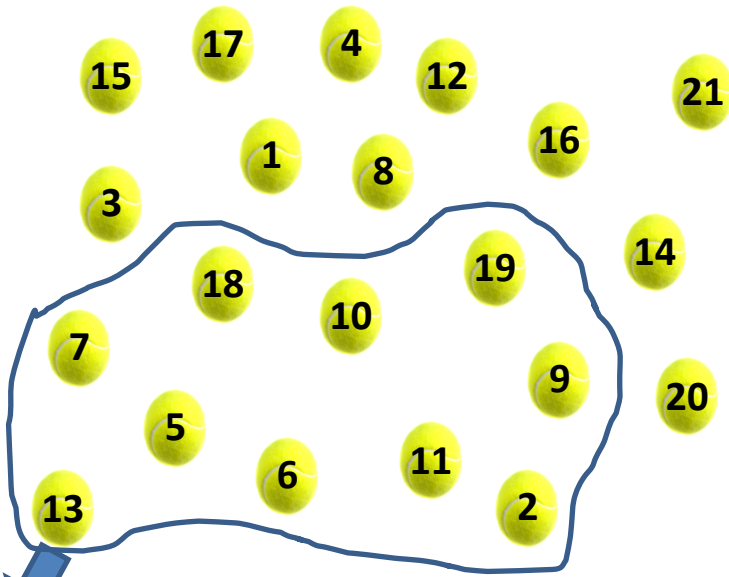
# Drawing the picture



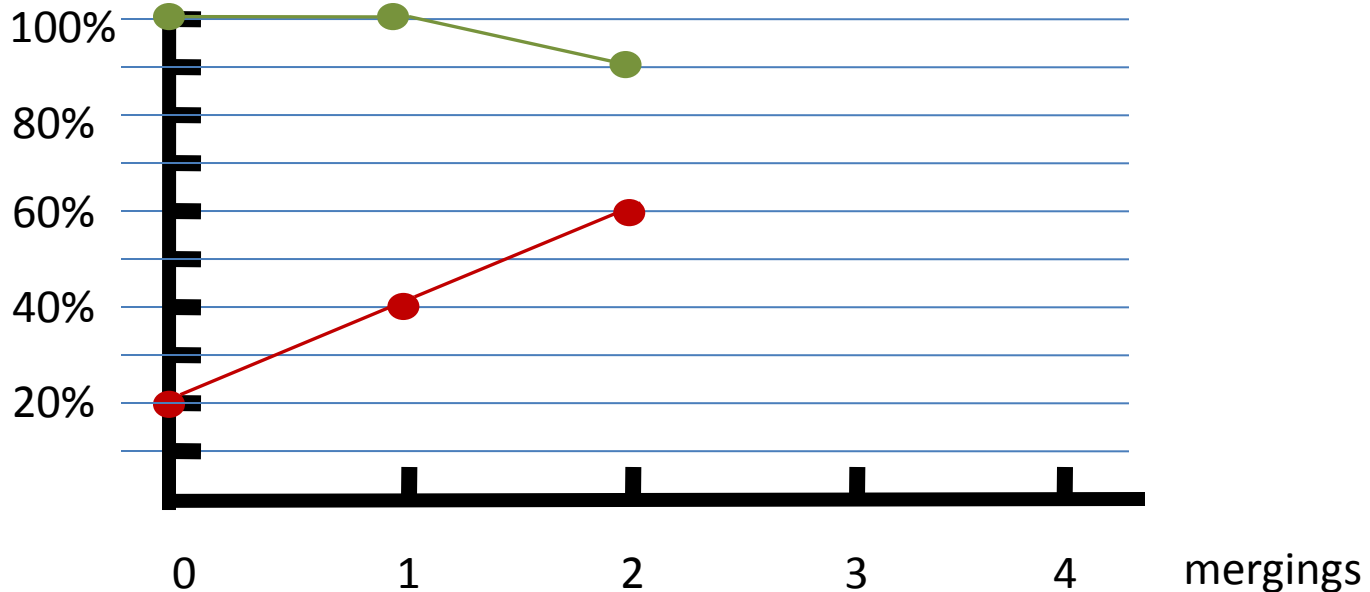
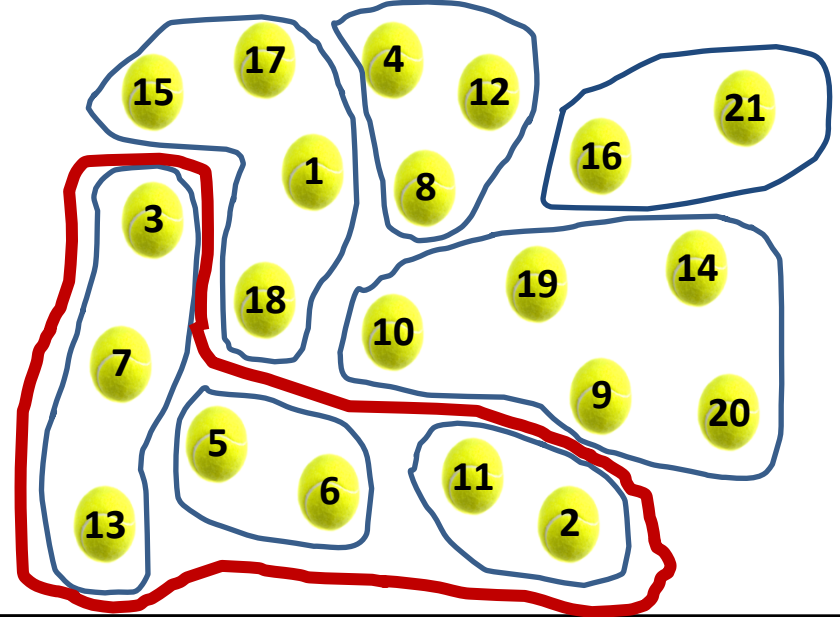
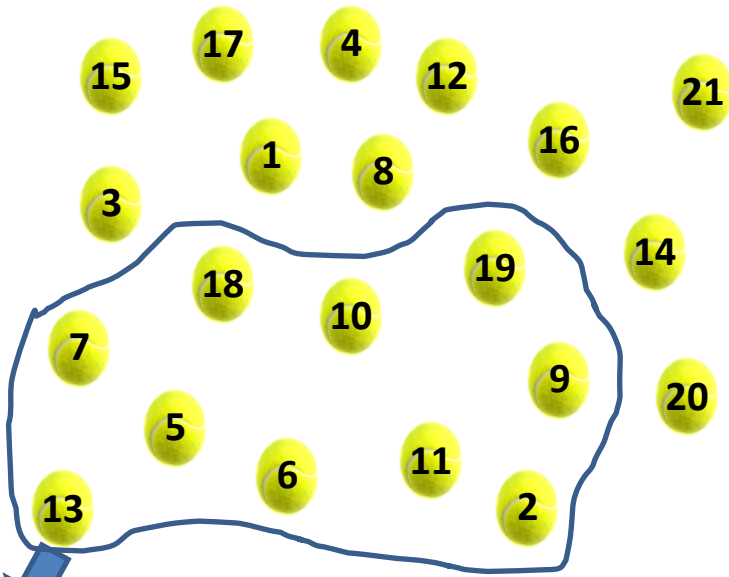
# Drawing the picture



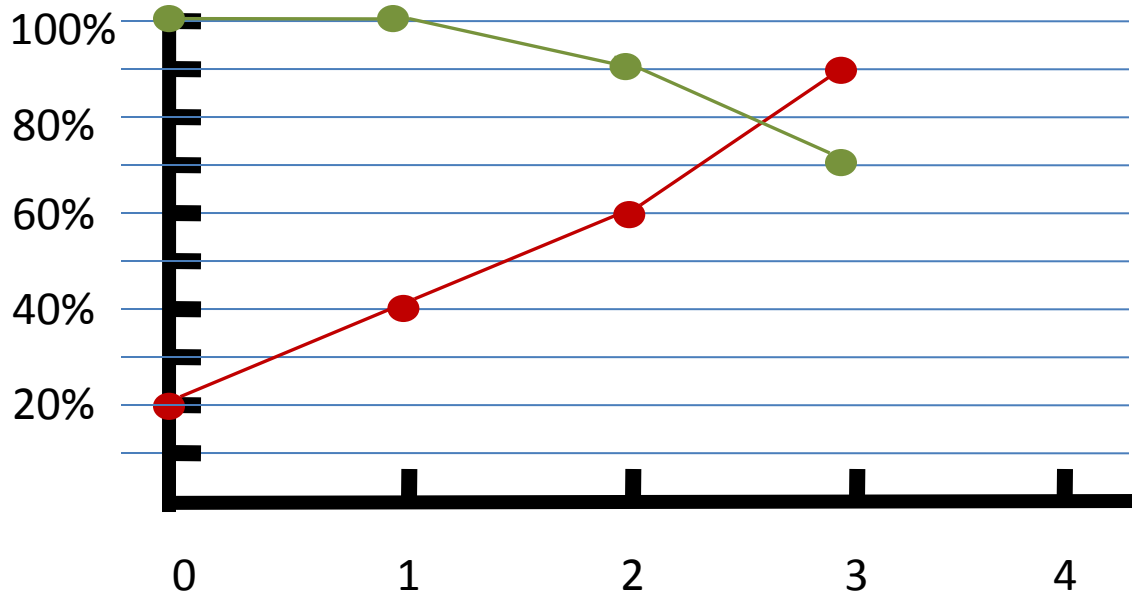
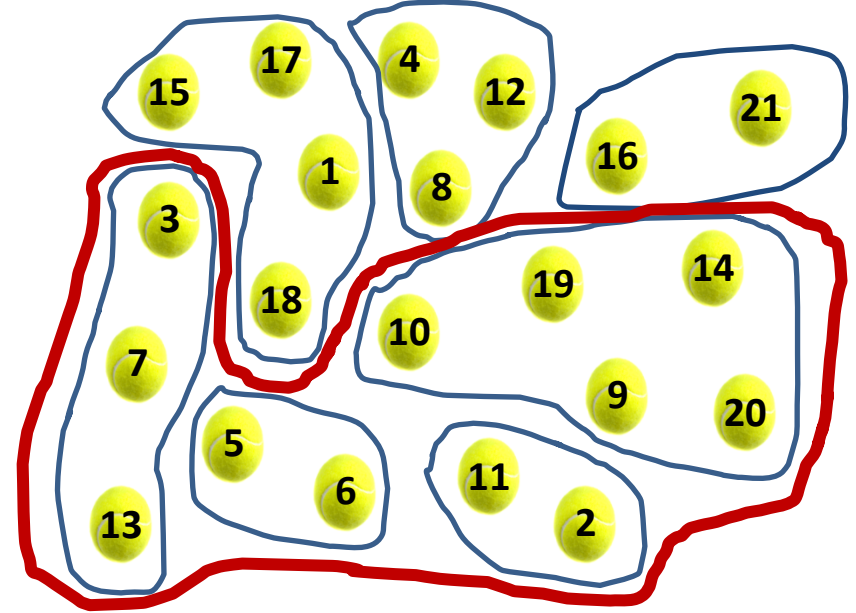
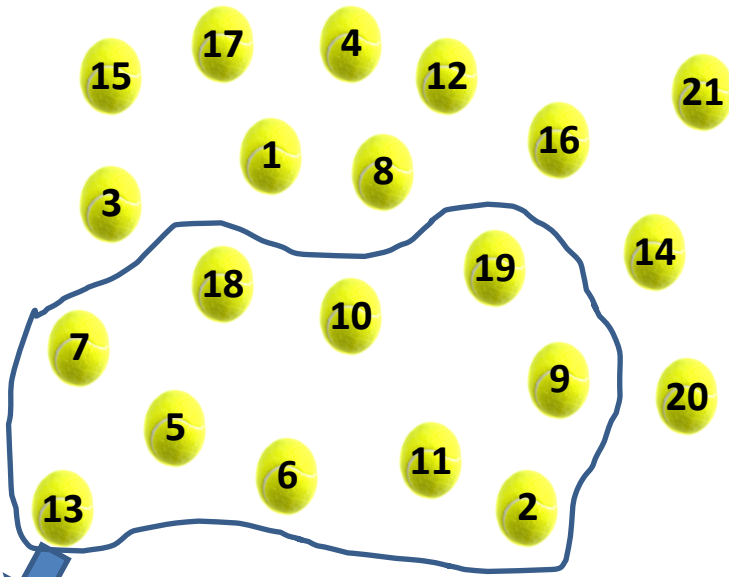
# Drawing the picture



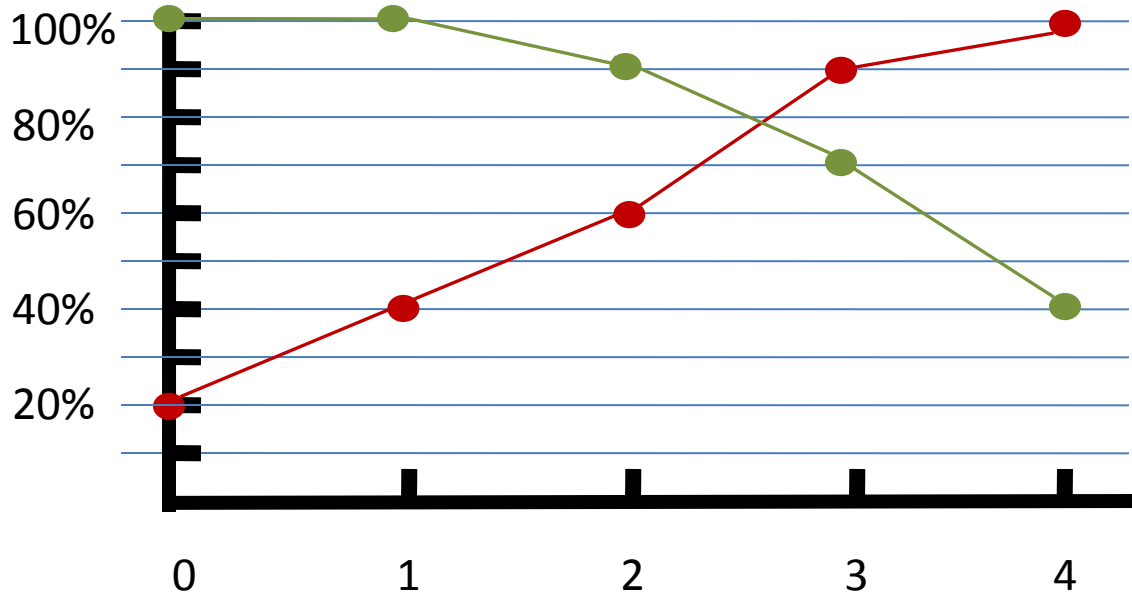
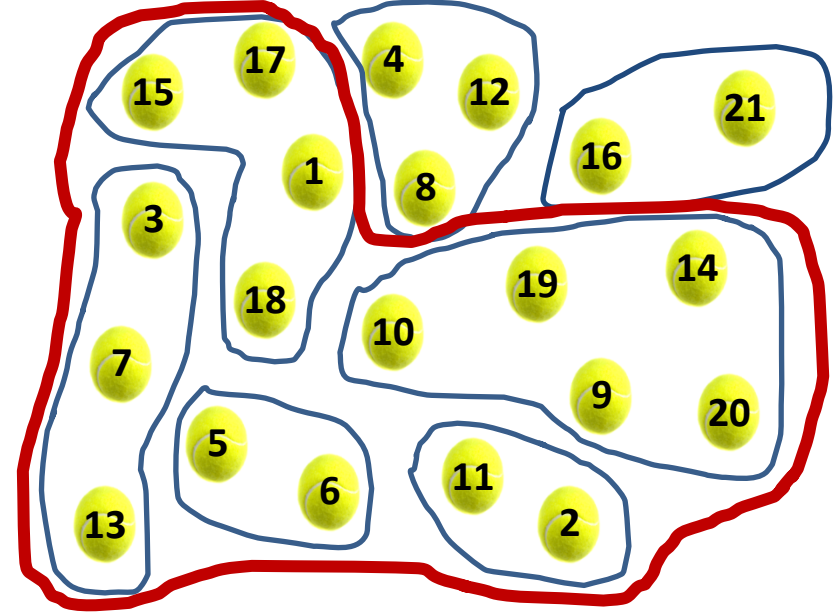
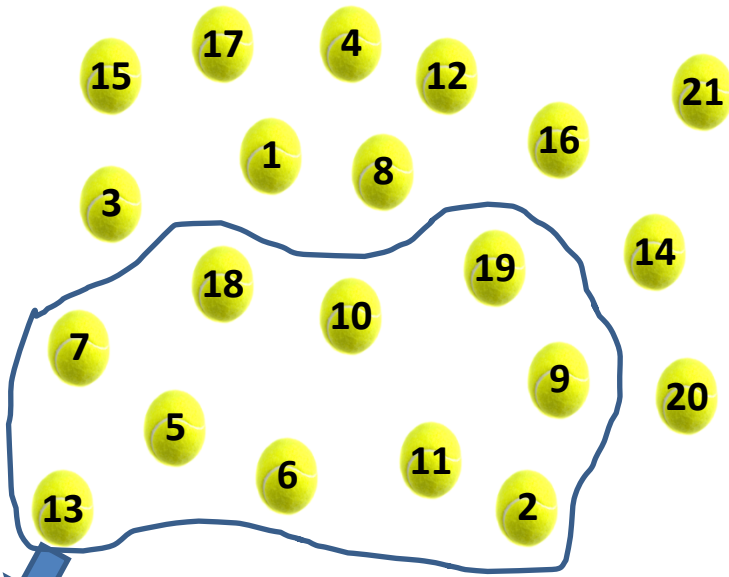
# Drawing the picture



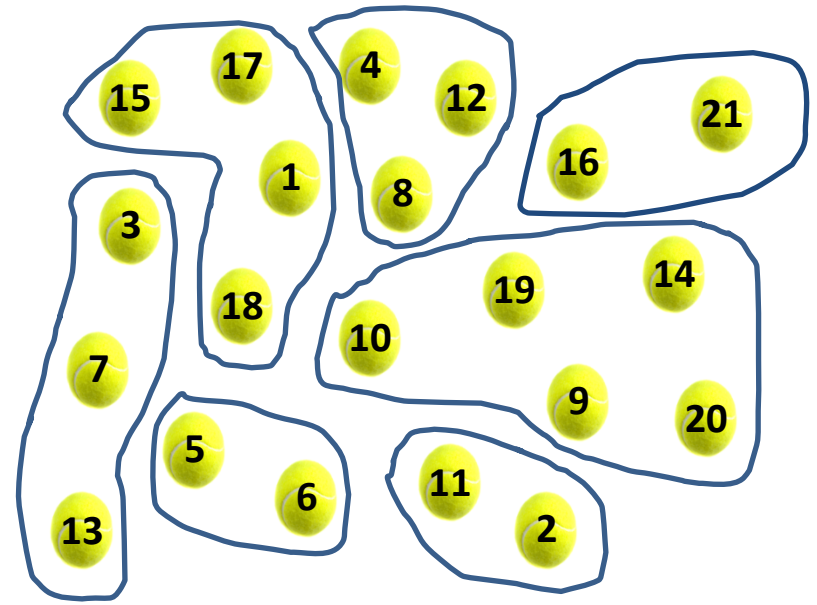
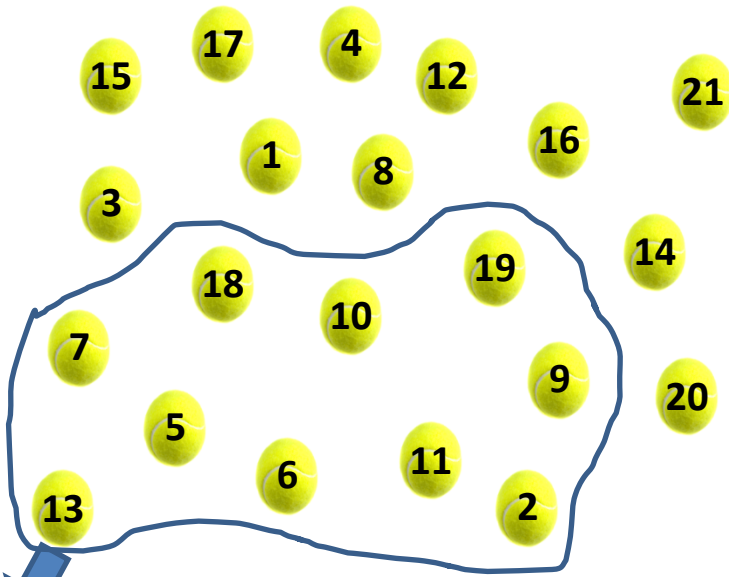
# Drawing the picture



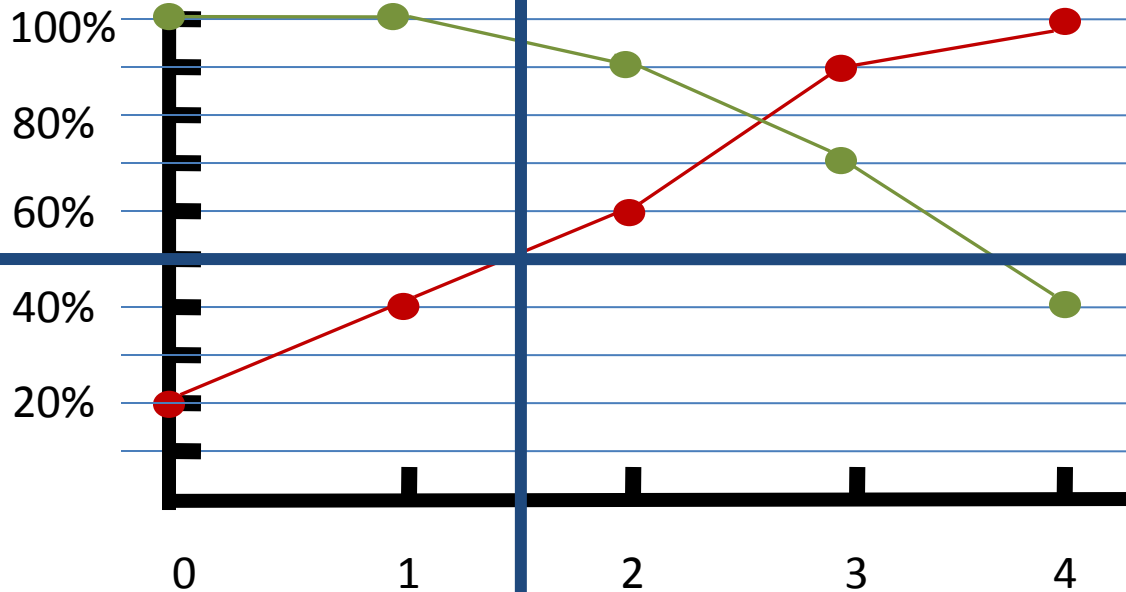
# Drawing the picture



# Drawing the picture

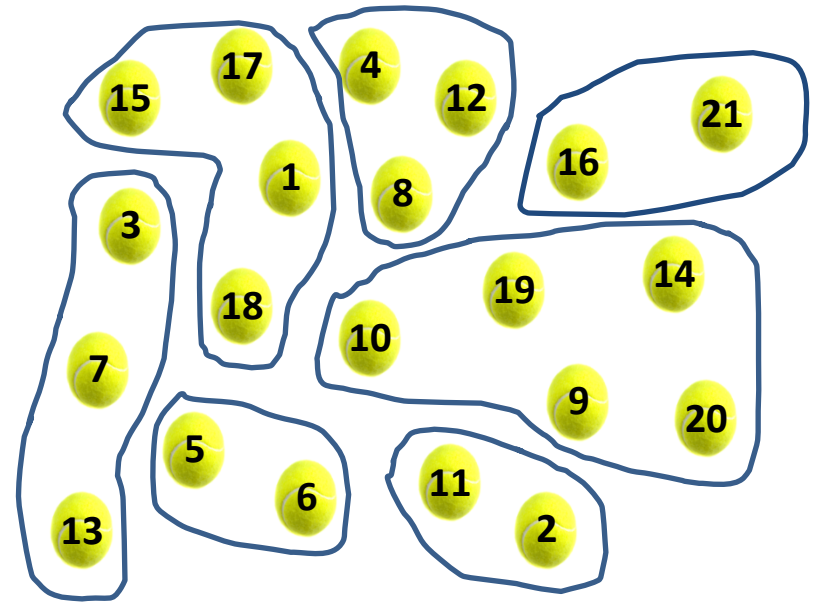
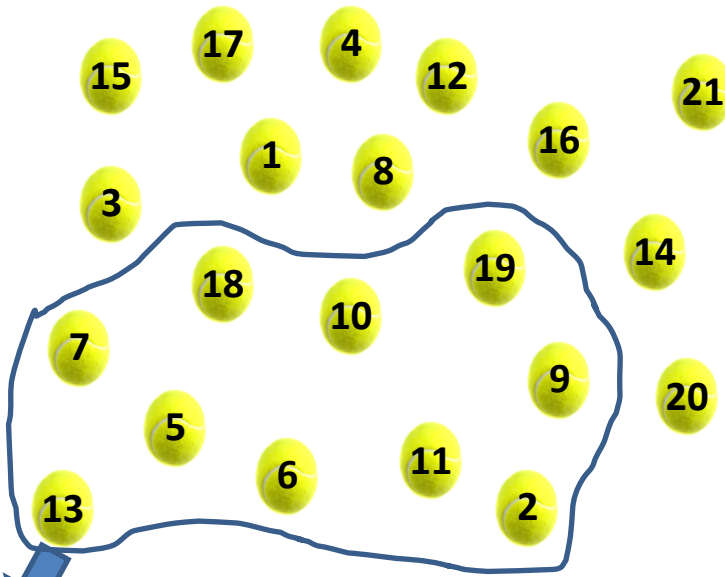


$t = 50\%$

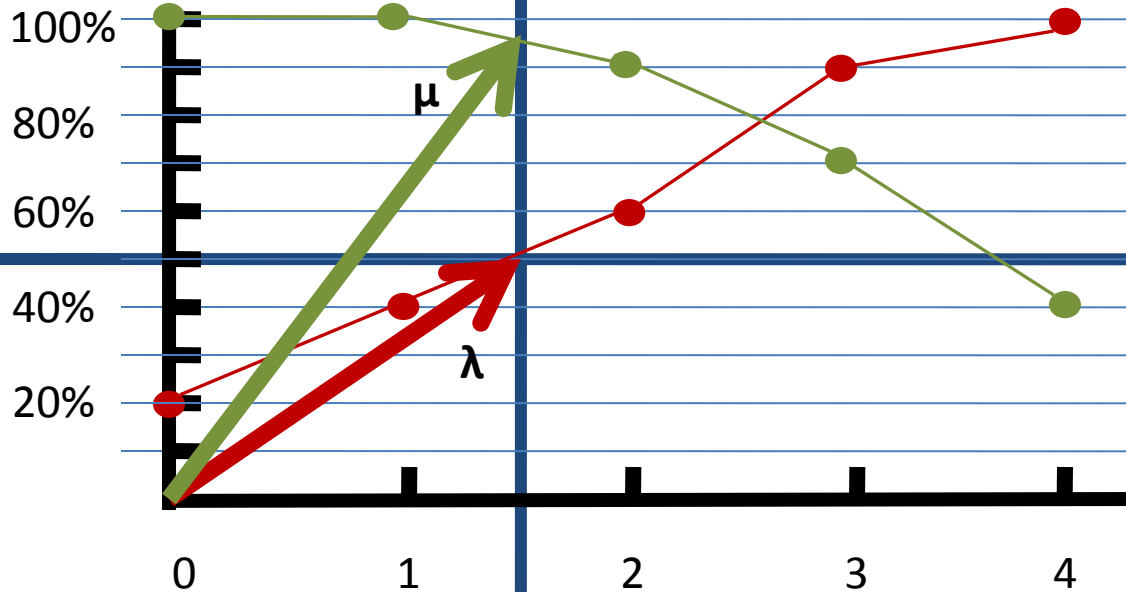


mergings

# Drawing the picture

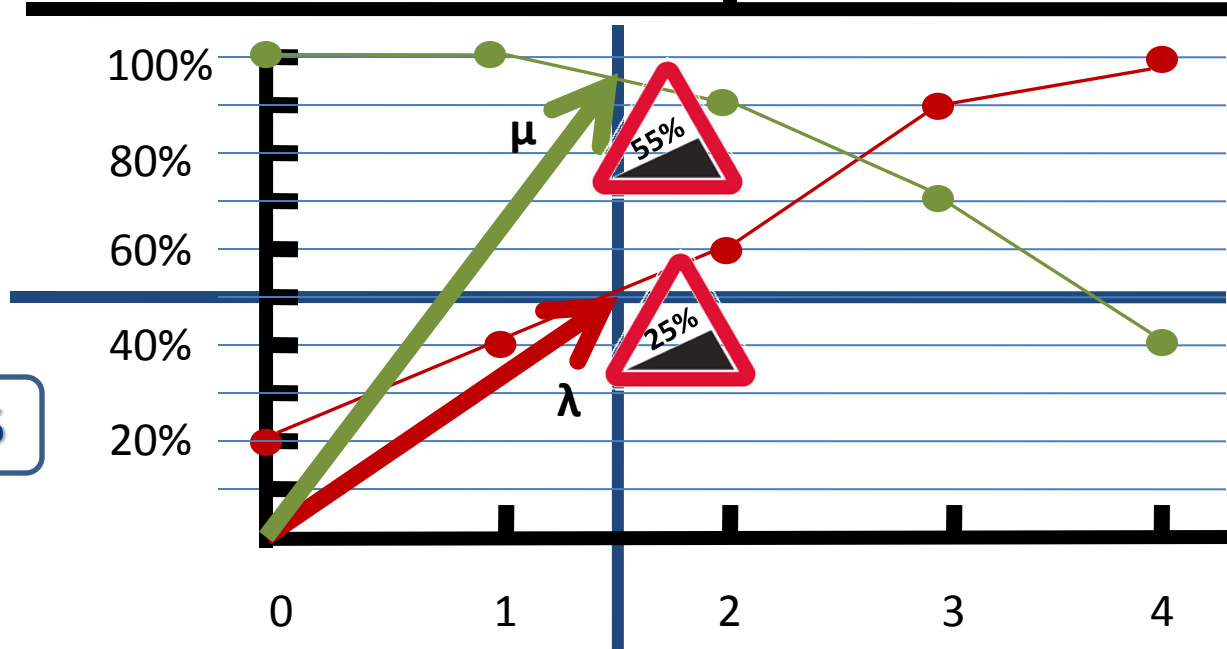
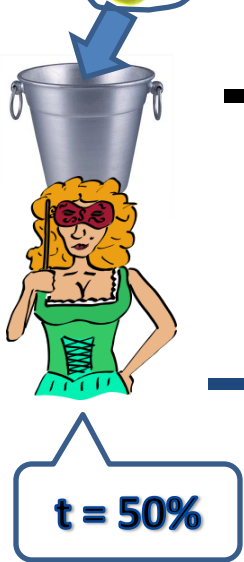
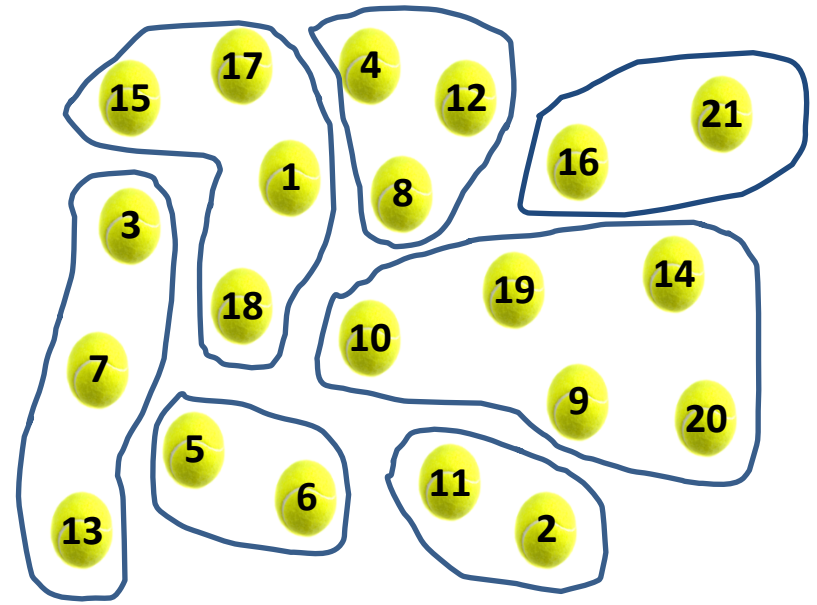
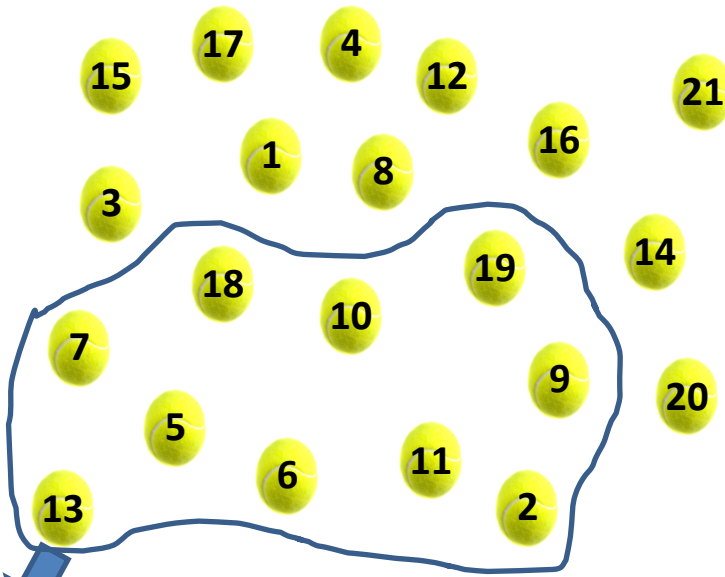


$t = 50\%$



mergings

# Drawing the picture

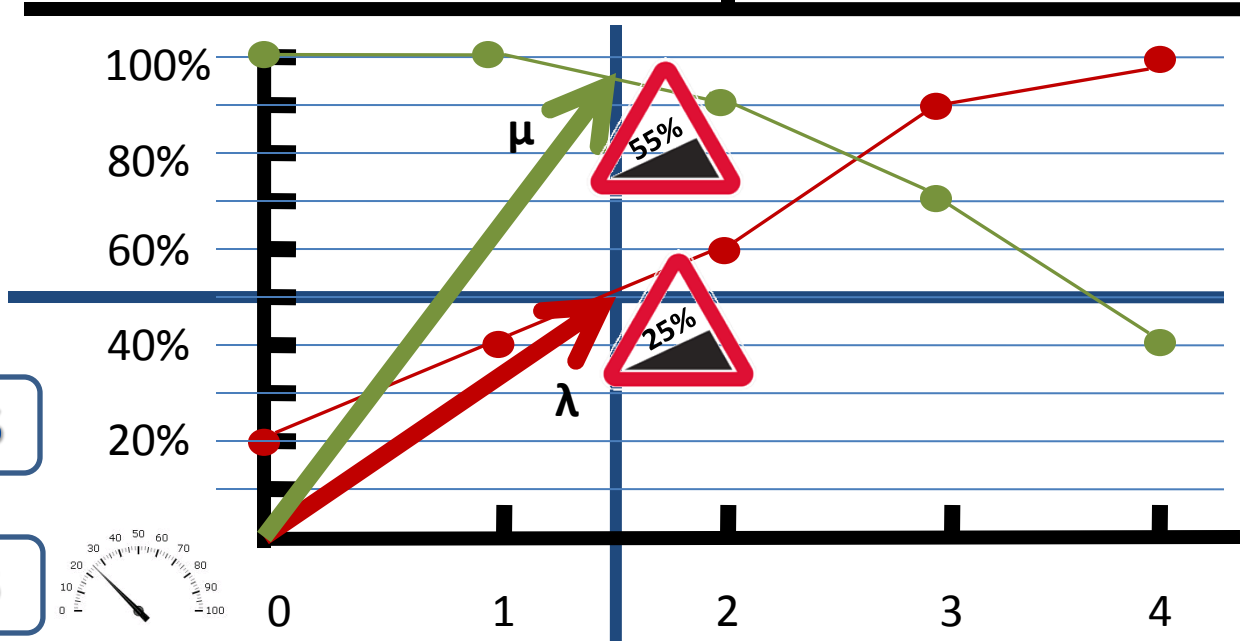
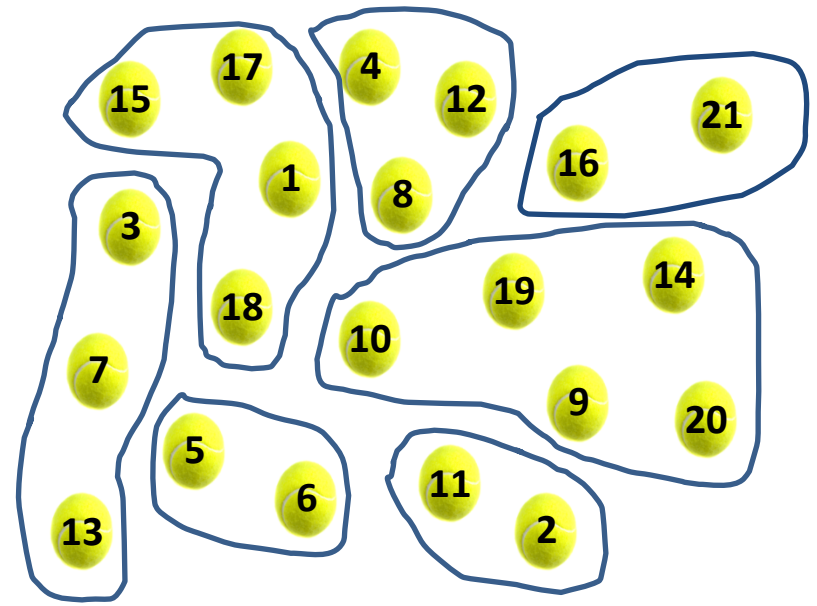
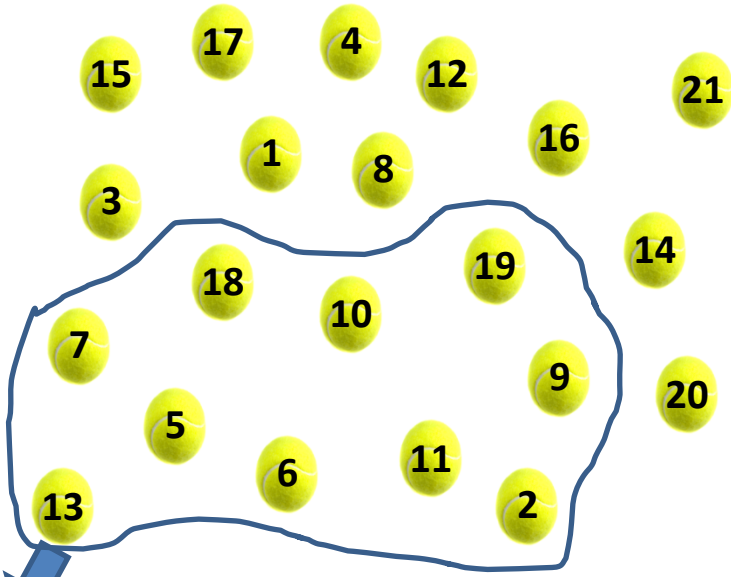


What is more important?

- Linking elements
- Maintaining purity

$$\text{RISK} = \alpha\lambda - (1-\alpha)\mu$$

# Drawing the picture



What is more important?

- Linking elements
- Maintaining purity

$$\text{RISK} = \alpha\lambda - (1-\alpha)\mu$$

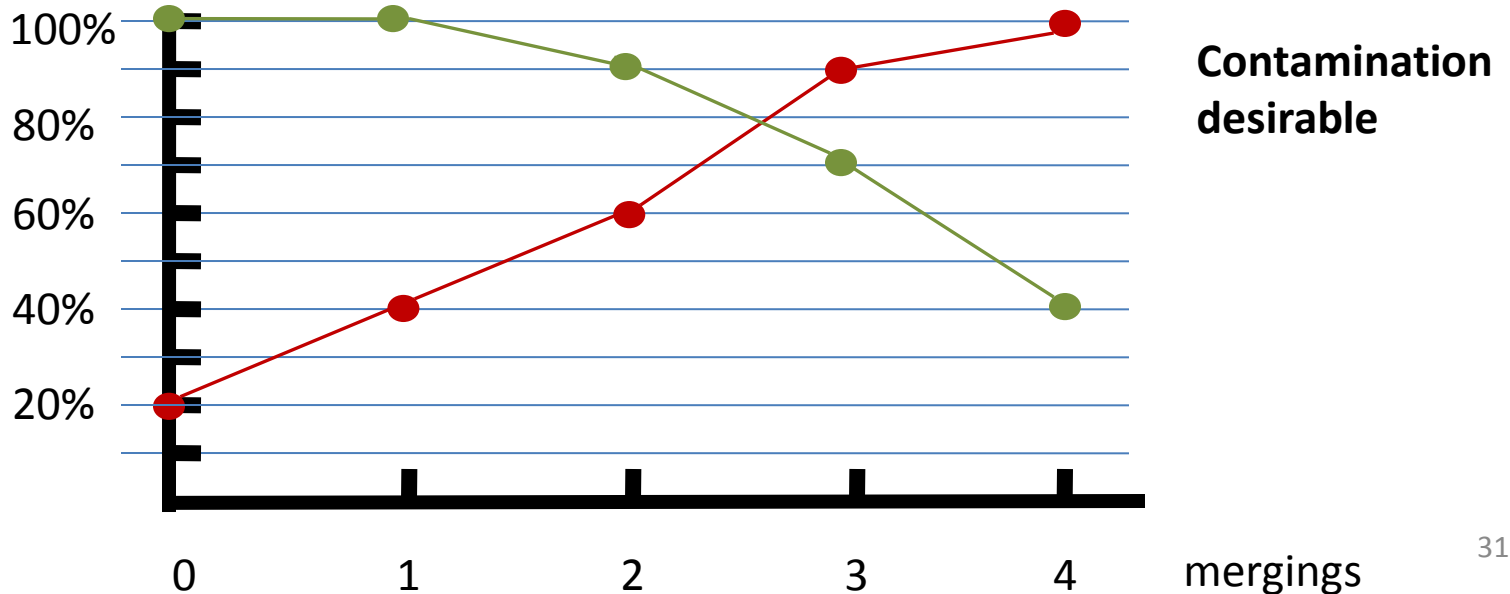
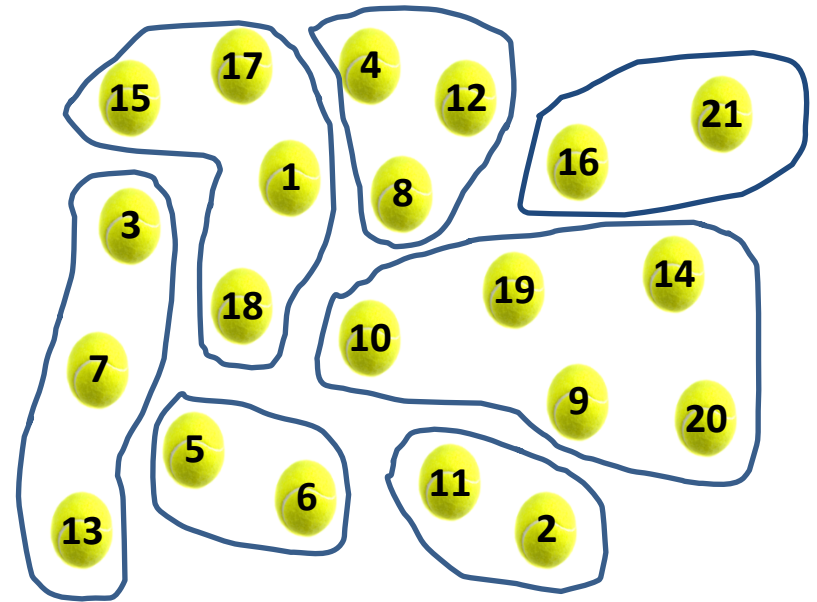
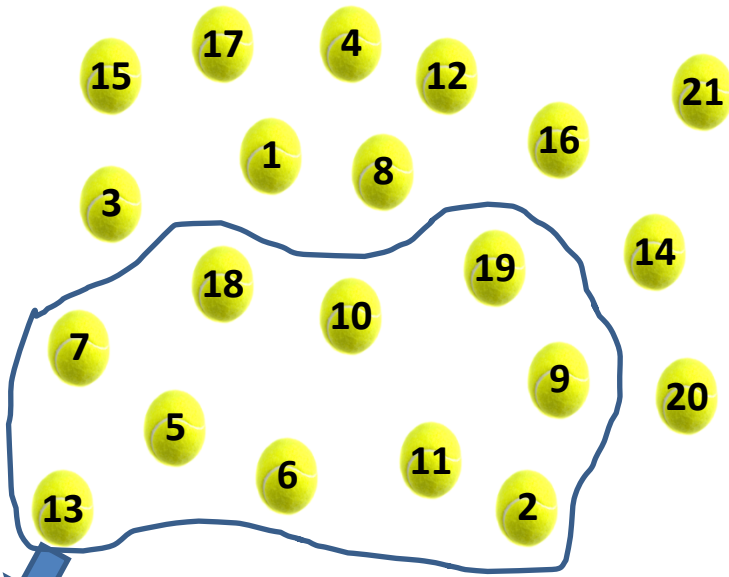
45%

mergings

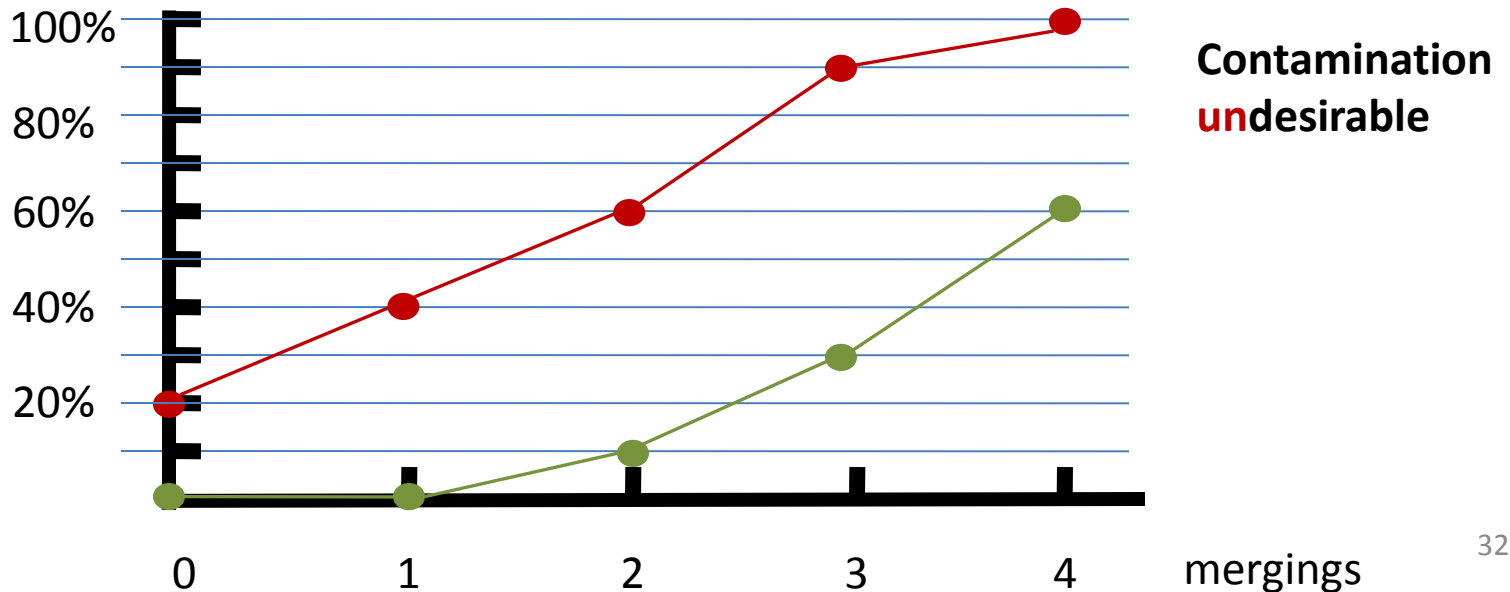
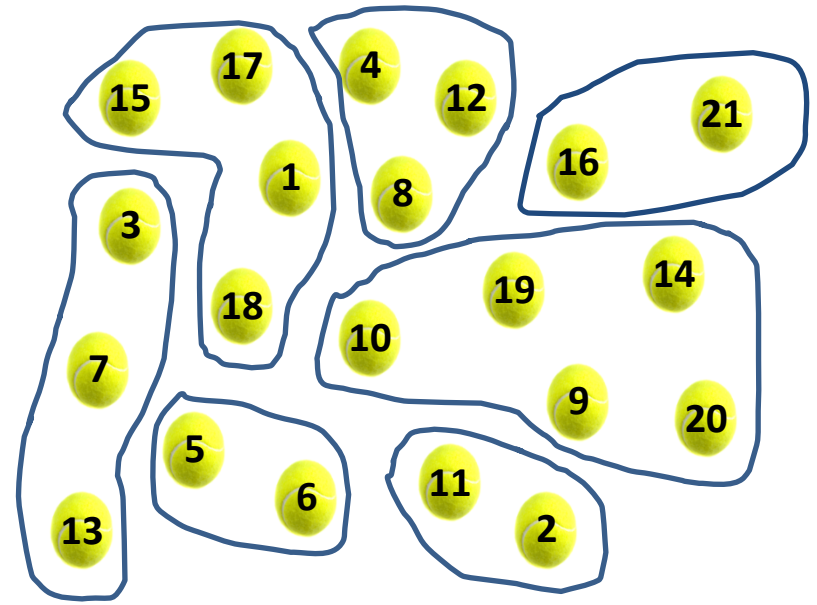
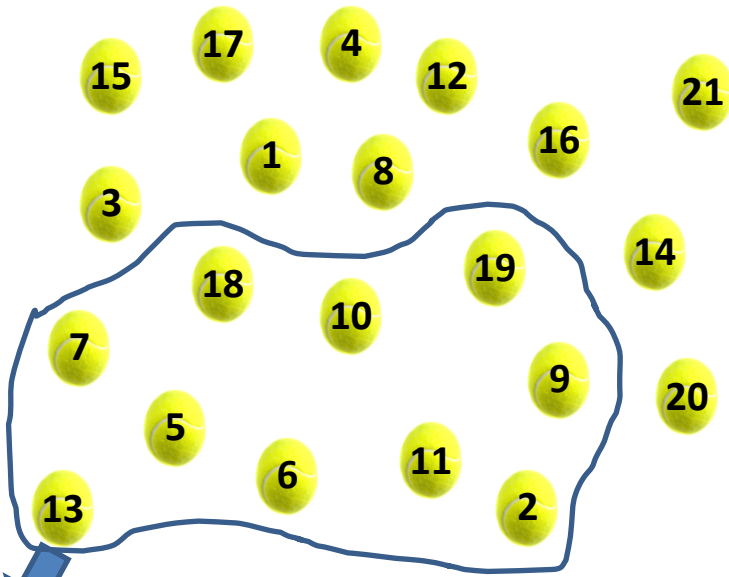
$t = 50\%$

$\alpha = 0.5$

# Drawing the picture

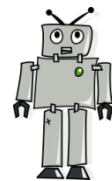
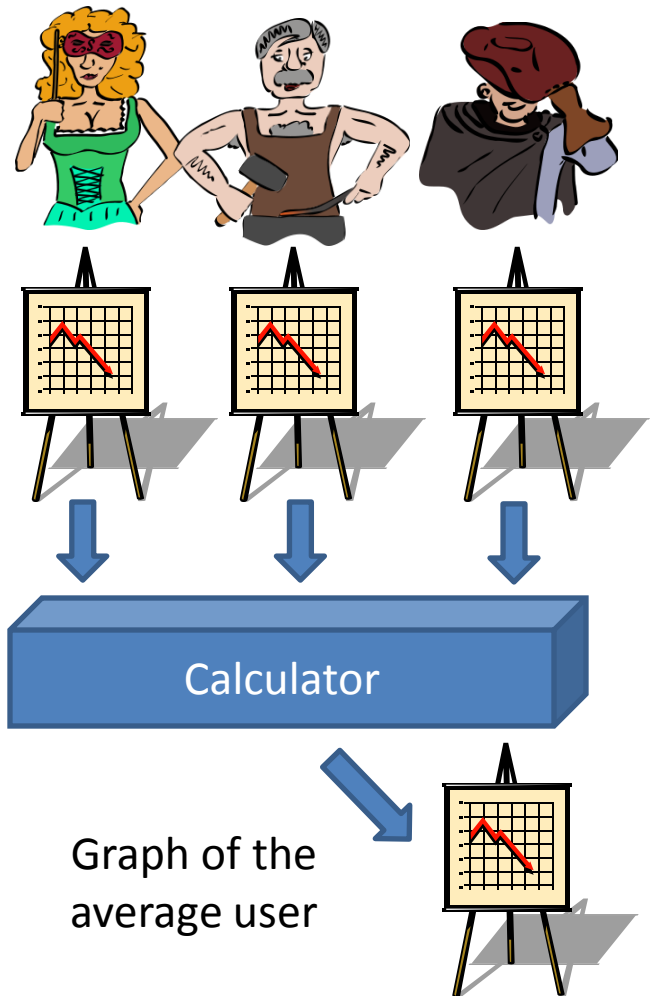


# Drawing the picture



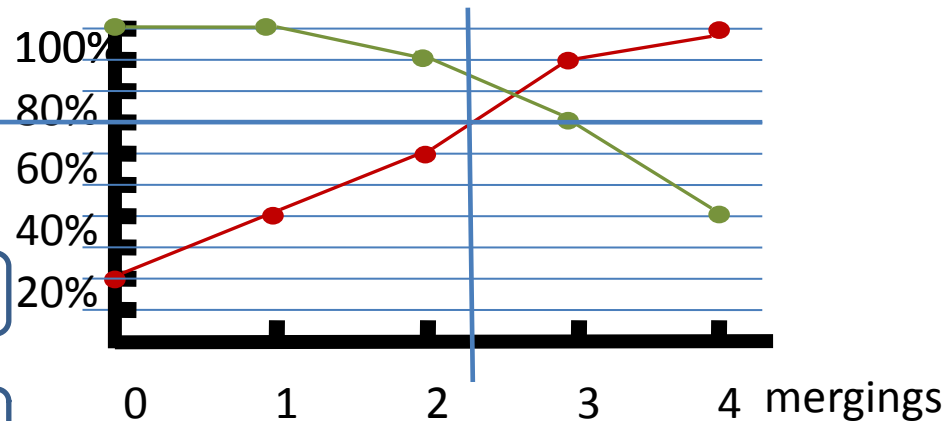
# Overall Linkability

- Overlays individual graphs
- Draws the average of each data point.
- Individual sensitivities are respected



$t = 70\%$

$\alpha = 0.5$



$$\text{RISK} = \alpha\lambda - (1-\alpha)\mu$$

# Outline

Motivation

The Idea

**Evaluation**

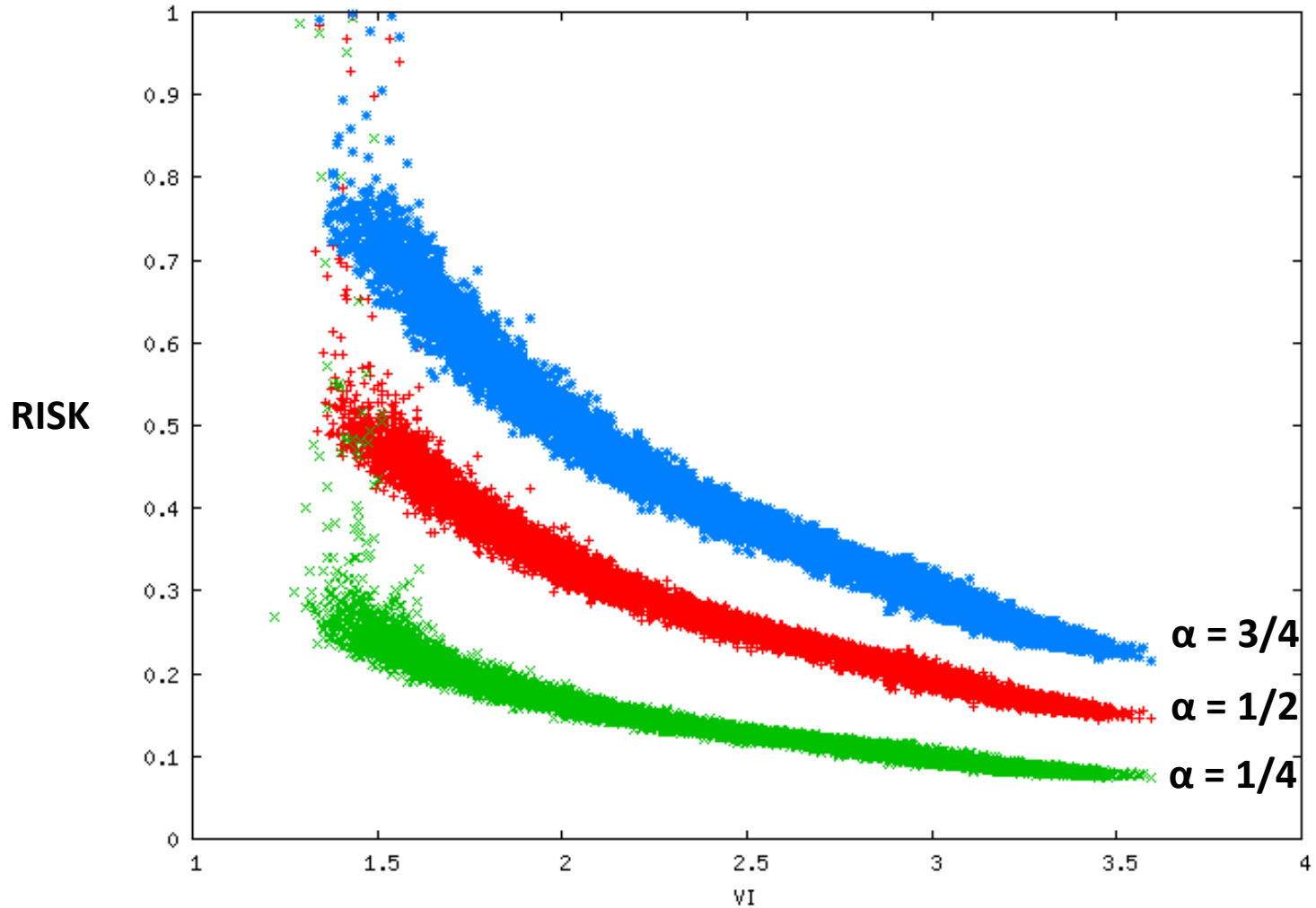
Conclusions

# Evaluation

- Does our overall “linkability risk” measure behave intuitively?
- We did the following experiment:
  - Generate random partition P of 200 elements
  - Start (many) random walks of different lengths from P;
  - Each random walk ends up in another partition
  - Plot the Variation of Information distance vs. our risk measure  $\Delta$

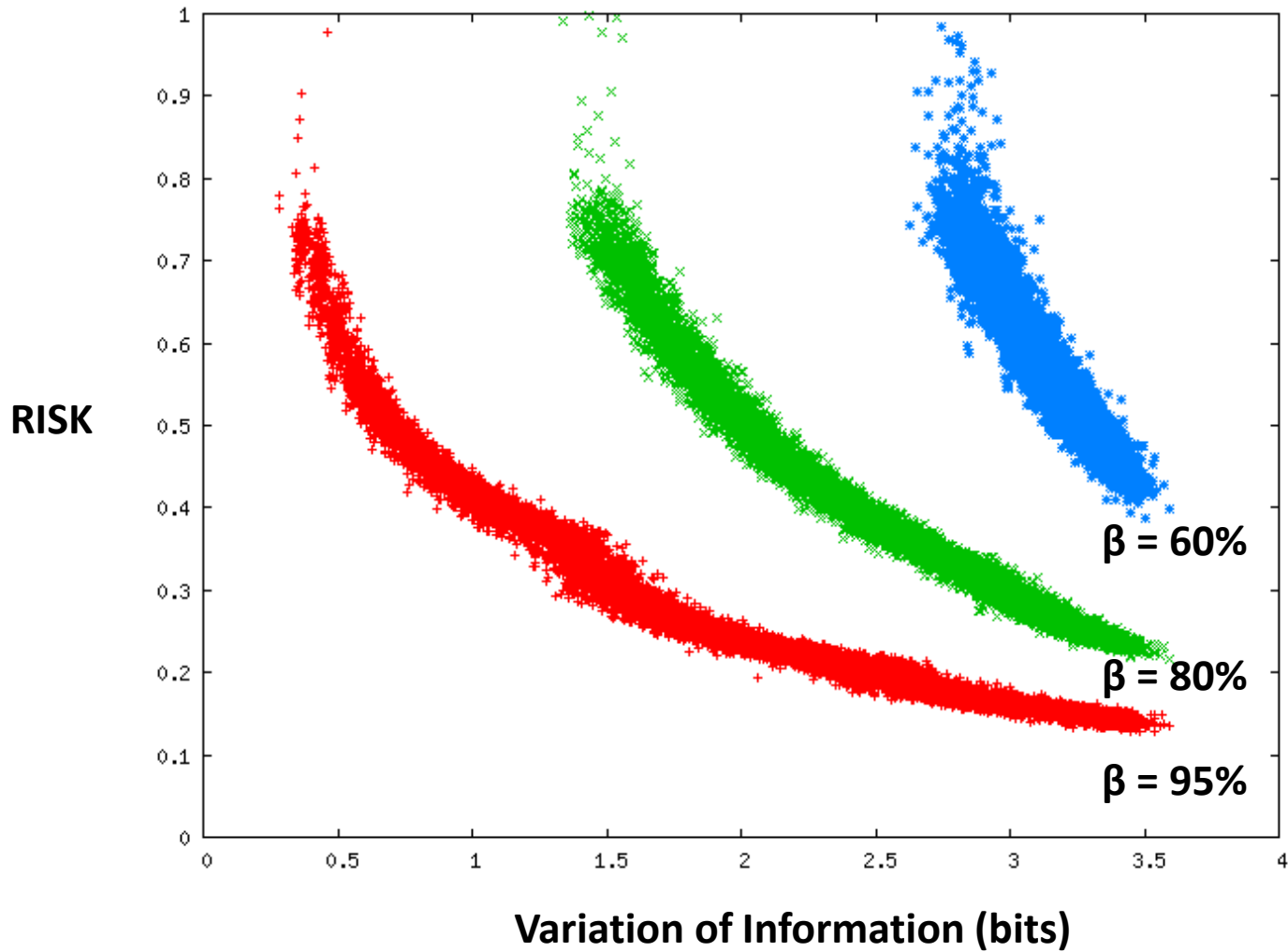
# Results

$n=200$ ; privacy threshold  $\beta = 80\%$ ;



# Results

$n=200;$     $\alpha = 3/4;$



# Outline

Motivation

The Idea

Evaluation

**Conclusions**

# Conclusions

- We introduced a ‘linkability risk’ measure that
  - operates on the subject level
  - lets evaluator specify
    - **sensitivity** of individual elements (or element types)
    - **importance** of linking vs. contamination (purity) effect
    - **privacy threshold**
  - is easily visualisable and hence intuitive
  - is consistent with real metrics over the solution space
  - provides natural ways to evaluate **fairness**

# Outlook

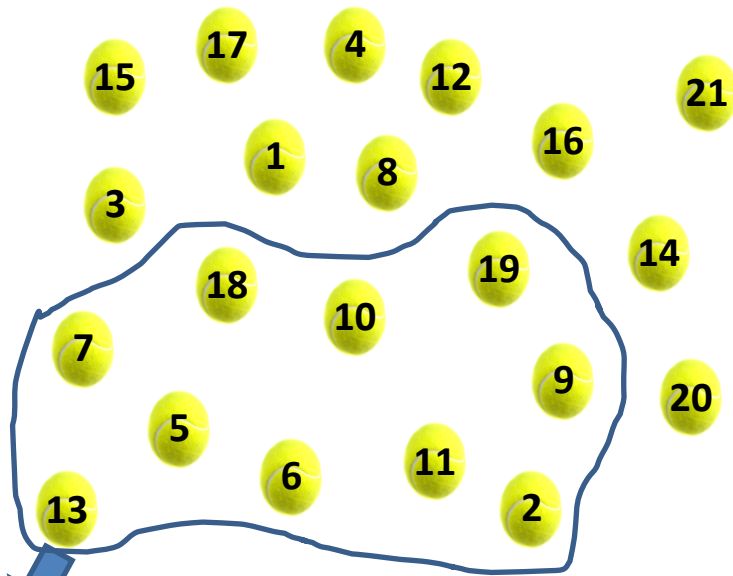
- Make measure more sensitive to ‘contamination quality’.
- Use our measure in practice, e.g. for evaluating attacks
- Can it be used to deliver feedback to end users?

**Thanks for listening!**

**Q<sub>s</sub>?**

# Backup Slides

# Drawing the picture



$m=9,$   
 $i=3, p=6$

$m=8, i=1,$   
 $p=4.5$

$m=7, i=2,$   
 $p=4.5$

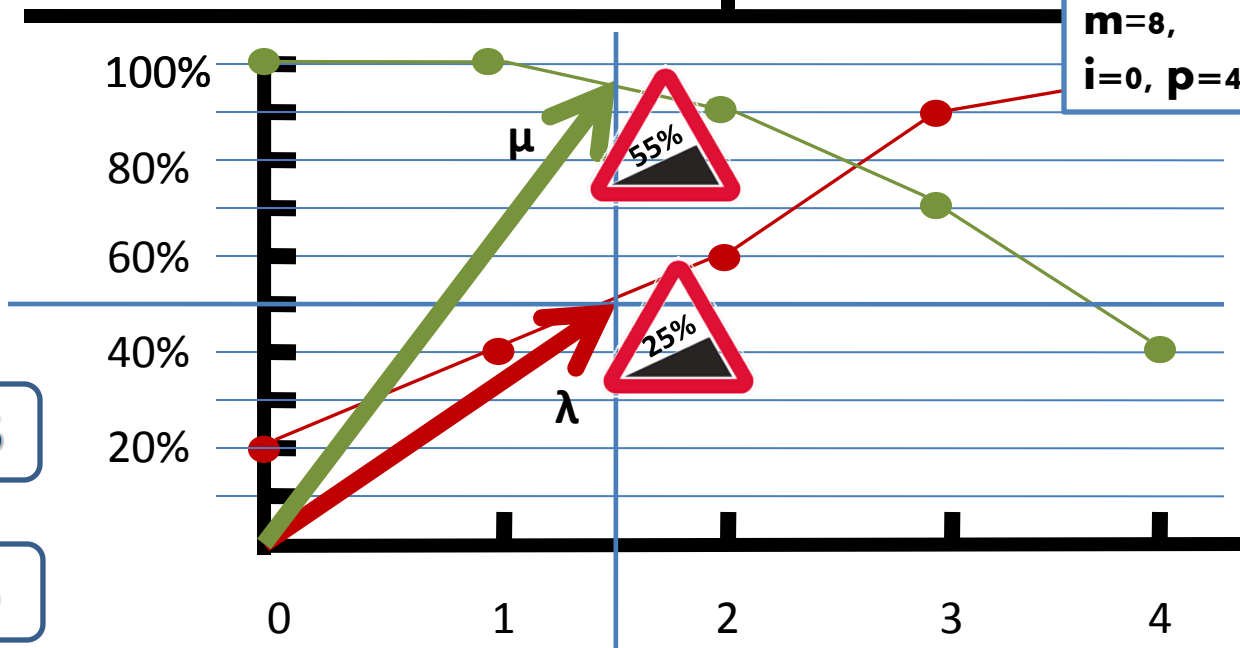
$m=8,$   
 $i=0, p=4$

$m=8,$   
 $i=0, p=4$



$t = 50\%$

$\alpha = 0.5$



What is more important?

- Linking elements
- Maintaining purity

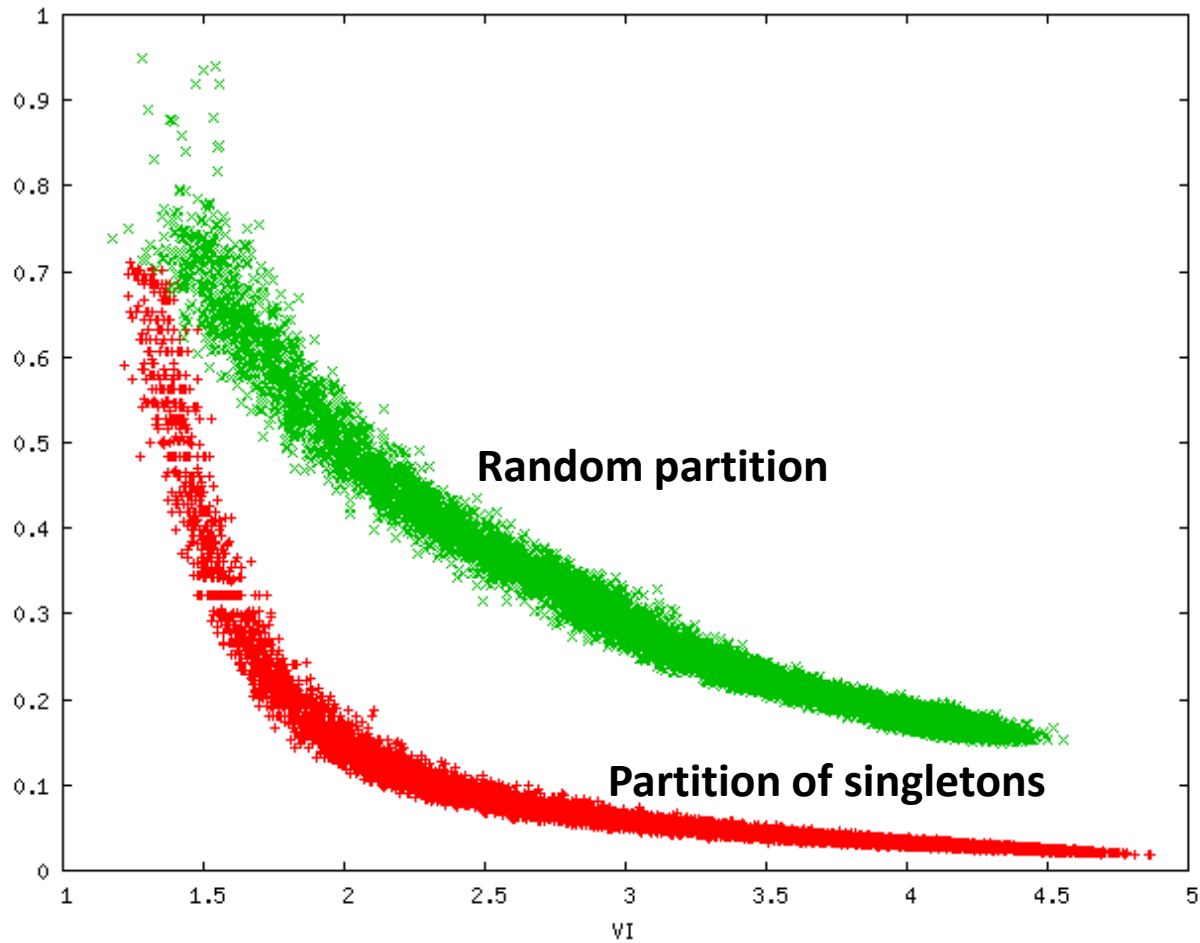
$$\text{RISK} = \alpha\lambda - (1-\alpha)\mu$$

45%

mergings

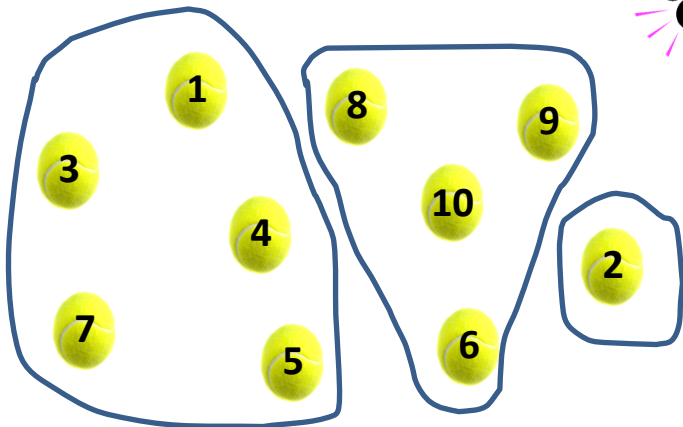
# Results

$n=200$ ;  $\alpha = 3/4$ ;  $\beta = 80\%$

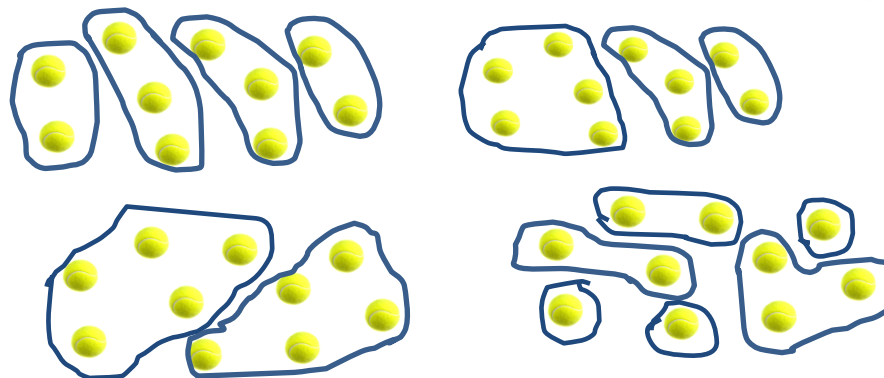


# Some unlinkability measures

$$P = \{P_1, P_2, P_3\}$$



View over solution space



Entropy of  
Adversary's  
view

[PET 2003]  
[PET 2007]

Per subject  
measurements

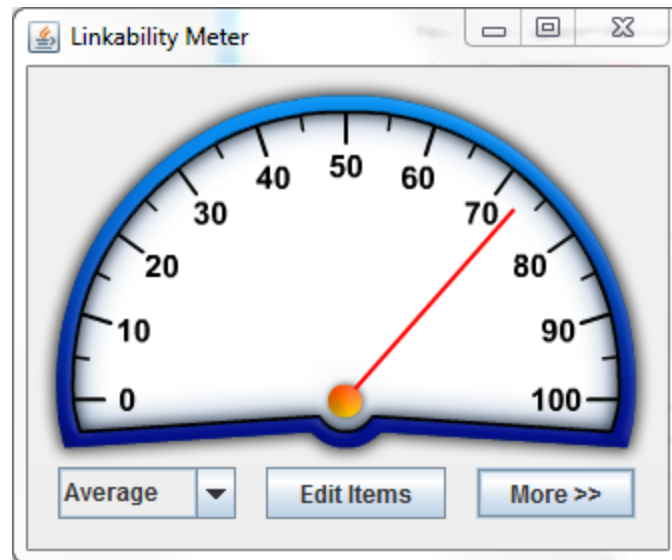
Enables fairness  
measurements

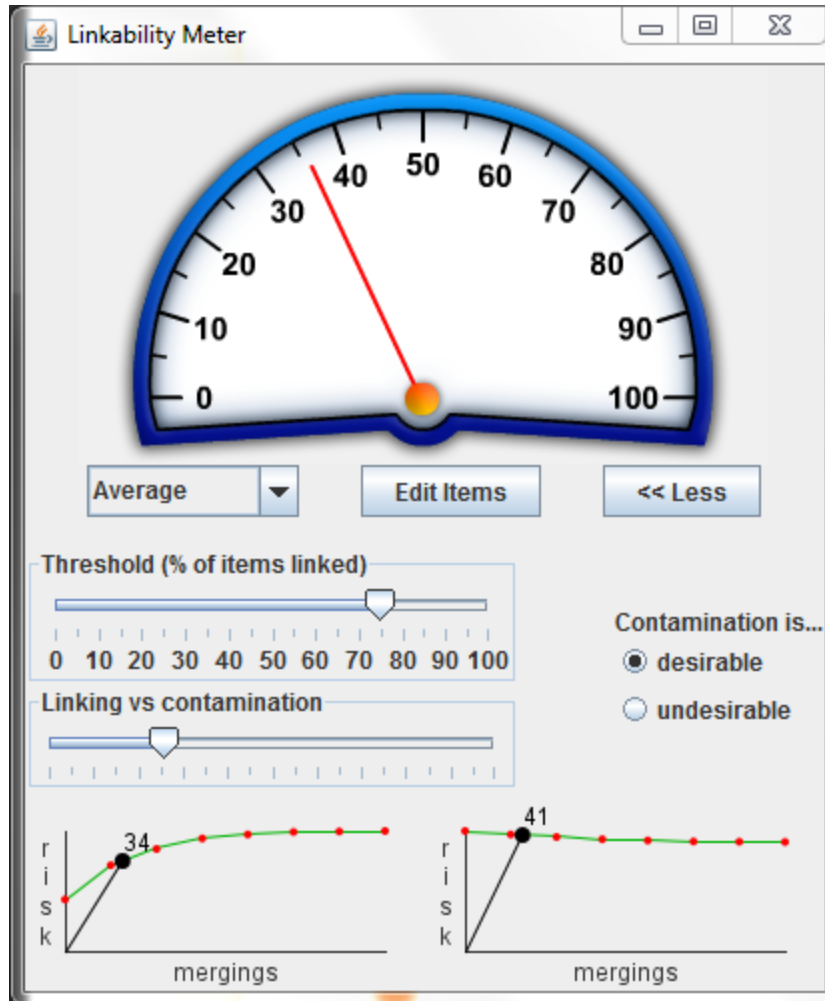
[APSCC 2008]

$$\sum_{P' \in \text{view}} \Pr(P') \delta(P, P')$$

$\delta$  is one of the  
distance measures

[WPES 2008]

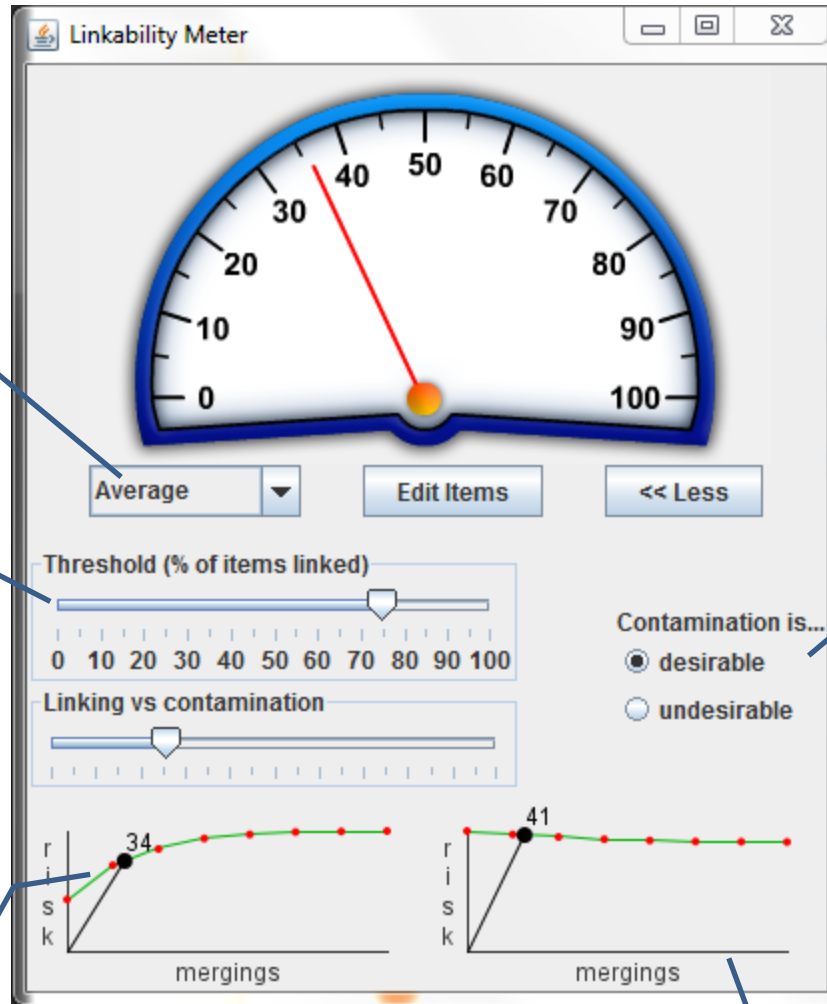




Select subject

Privacy threshold

Linking Graph



Mode

Contamination Graph