Entropy algorithm

NHA NGUYEN - HMCLAB

PROF. MINHO SHIN

Location entropy of 1 user

- Given a user u and its cloaked area R,
  - R is partitioned into m cells.
- S(R) = {c_1, c_2, ..., c_m}: set of cells inside R
- C_U: the cell that contains user U
  - random variable with possible values {c_1, c_2, ..., c_m}
- The entropy of C_U is

$$\left\lceil \frac{x_u - x_{\min} - d}{x_{\max} - x_{\min} - d} p \right\rceil, \left\lceil \frac{y_u - y_{\min}}{y_{\max} - y_{\min}} q \right\rceil$$

where P_i is the probability user U is located in c_i

$$P_i = \Pr(C = c_i) = \{\text{Prob} : l_u \in c_i, c_i \in S(R)\}$$

  - l_u: location of user u

Location entropy of 1 user

- Let R_1, R_2, ..., and R_n be the stream of cloaked areas of user U generated at time t_1, t_2, ..., and t_n.
  - where $t_1 < t_2 \cdots < t_n$
- $C(t_j)$: the cell that contains user U at time t_j
- The probability user U is located in c_i

$$P_i = \frac{\sum_{j=1}^{n} [C_U(t_j) = c_i]}{n}$$

Location entropy of 1 user

$$0 \leq H(C) \leq \log m$$

- H(C) = 0 : There is a cell that has 100% chance to be the cell that contain of user. i.e., the adversary can identify exactly the cell that the user is located.
- H(C) = log *m* : every cells in R has an equal chance to be the cell that contain of user

Location entropy of k users

- Given k users $\{u_1, u_2, ..., u_k\}$ in the simulation
- Because the cells that contains users in a cloaked region are mutually independent, entropy:

$$H = H(C_1, C_2, ...C_k) = \sum_{l=1}^{k} H(C_l) = -\sum_{l=1}^{k}\sum_{i=1}^{m} P_i \log P_i$$

$$\text{where } P_i = \frac{\sum_{j=1}^{n} [C_l(t_j) = c_i]}{n}$$

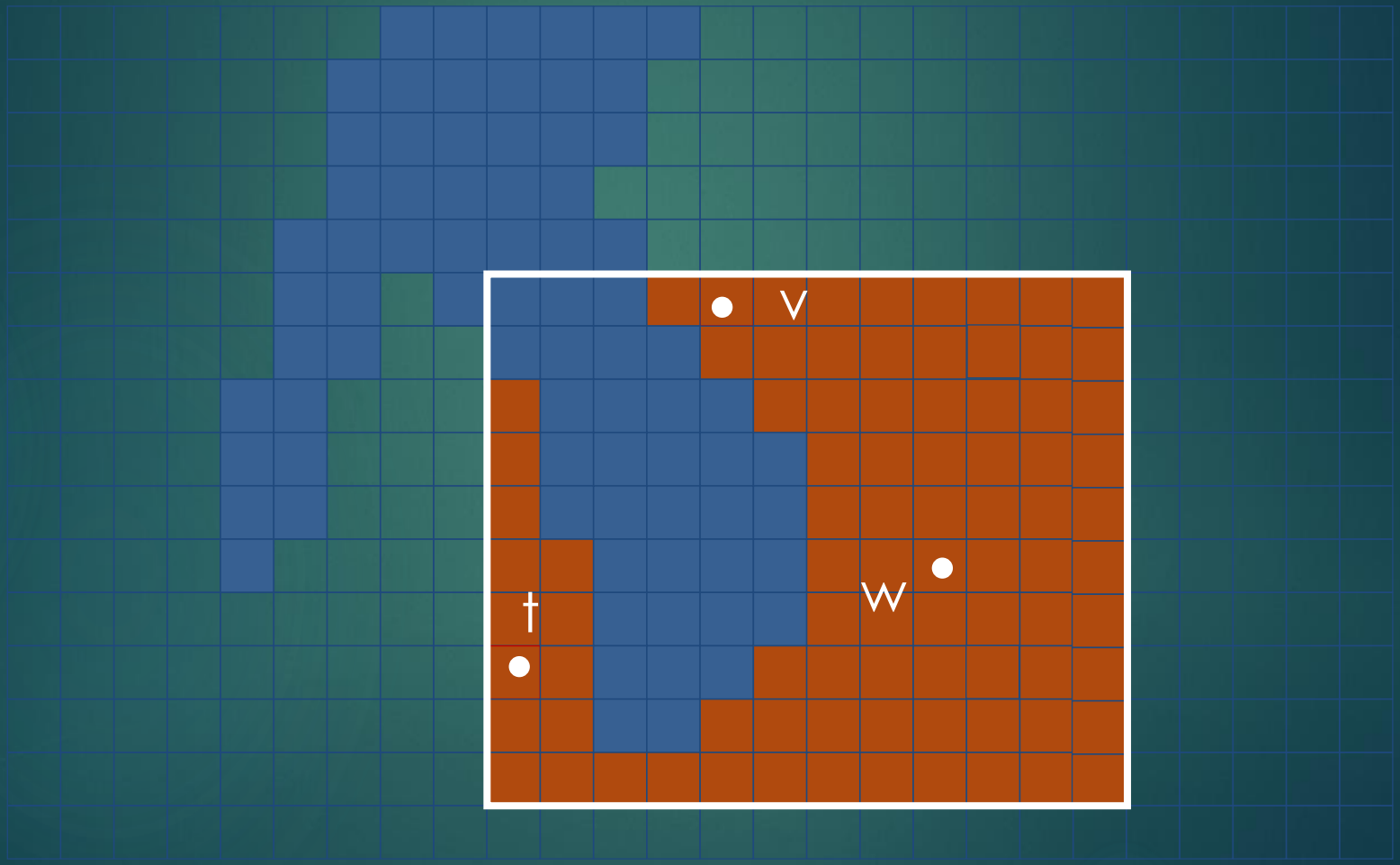
- $C_l(t_j)$: the cell that contains user u_l at time t_j

Algorithm

- **Input:** Location of each user,
- **Output:** Entropy

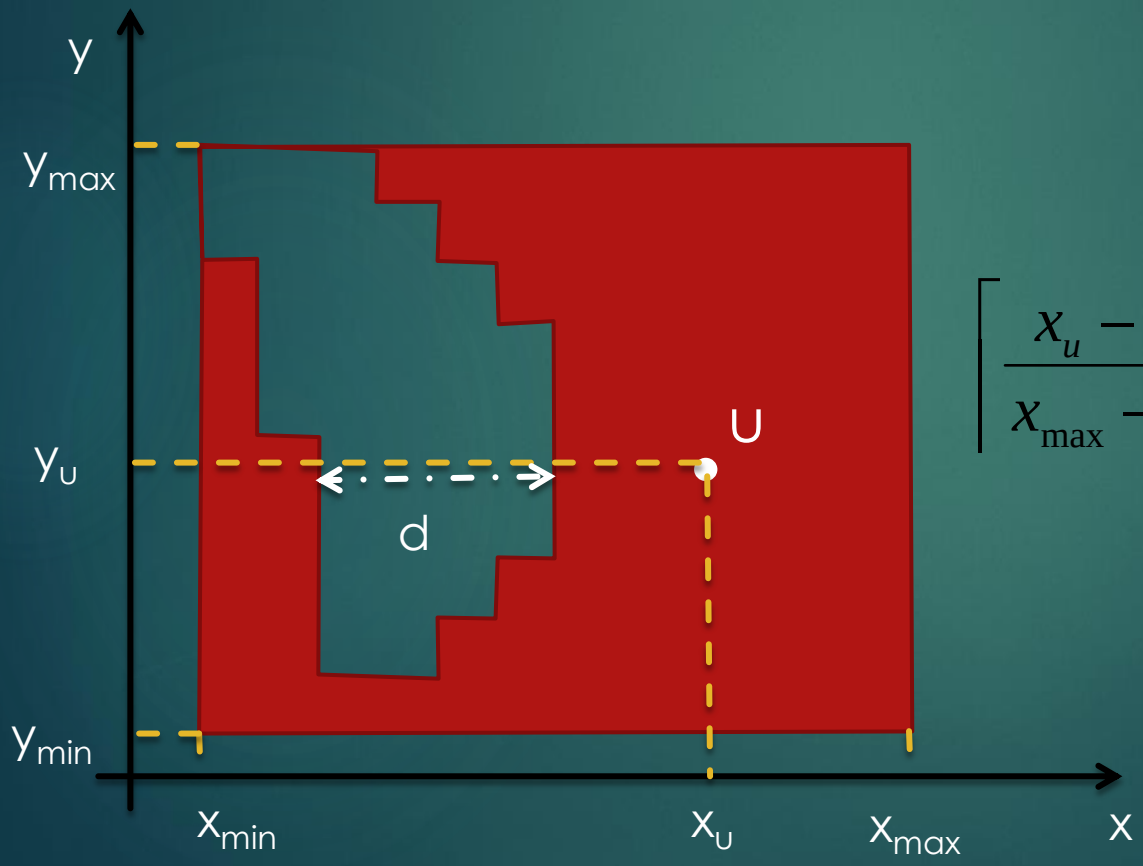
1. **for** each user u_1 **do**
2. **for** each cell c_i **do**
3. **for** each cloaked region R_j at time t_j $(1 \le j \le n)$ **do**
4. **if** u_1 is located in cell c_i **then**
5. count ++;
6. $P_i \leftarrow$ count/n;
7. $H_1 \leftarrow -\sum P_i \log P_i$
8. $H \leftarrow \sum H_1$

Define which cell contains user U



Define which cell contains user U

- R is partitioned into m = p x q cells
- Coordinate of user U: x_u, y_u



- Index of cell that contains user U

$$\left\lceil \frac{x_u - x_{\min} - d}{x_{\max} - x_{\min} - d} p \right\rceil, \left\lceil \frac{y_u - y_{\min}}{y_{\max} - y_{\min}} q \right\rceil$$