Detecting Aged Deterioration of a Radio Base Station Map for Wi-Fi Positioning

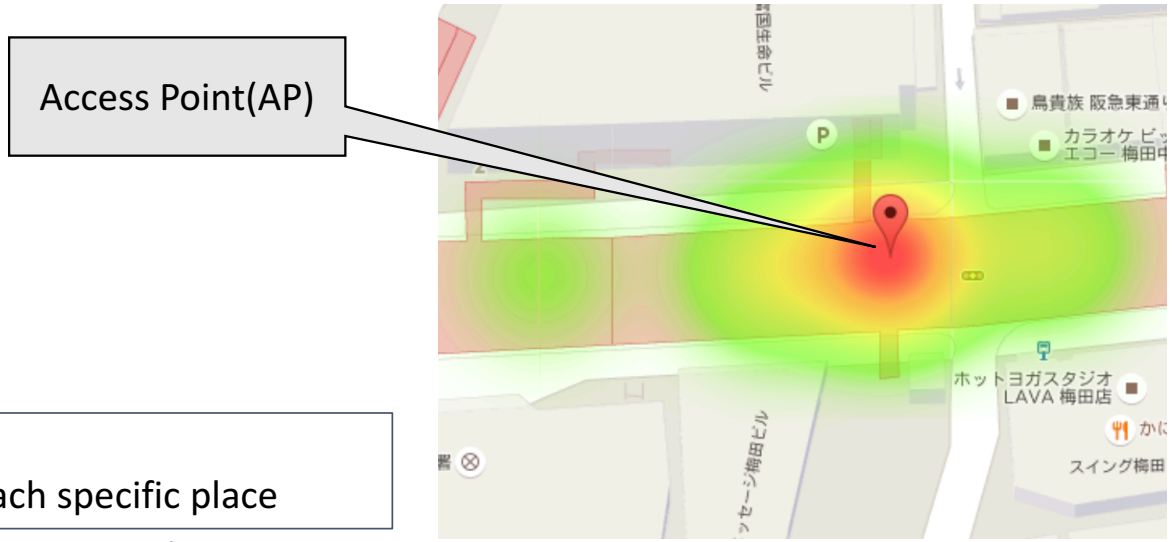
Sep. 11, 2017

Makiko Kawanaka[†1], Kohei Yamamoto[†1], Kota Tsubouchi[†2],
Kazuya Murao[†1], Masahiro Mochizuki[†1], Nobuhiko Nishio[†1]

[†1]: Ritsumeikan University, [†2]: Yahoo Japan Corporation

Background

- Indoor positioning is investigated at where GPS cannot reach
- Wi-Fi localization utilizing already-installed AP in public area is focused
- Localization model must be created beforehand



Localization model：
A model reflects Wi-Fi signal strength at each specific place

Problems of Wi-Fi localization

- Problem 1:
  The accuracy of model fidelity decreases due to environmental change including appearance, disappearance, and relocation of AP
  - ➢ When and where of a model to be updated is unknown although it needs to be done regularly
  - ➢ Updating cost is high since it has to be renewed including all the AP
- Problem 2:
  Selection of AP thought to be stable under operation used for localization is limited and heuristic
    - e.g.) AP showing ESSID with career's name are thought to be stable
  - ➢ Stability of AP are not evaluated
  - ➢ Stable AP must be selected and used for localization

Research Objective

- To detect AP supposed to be updated
  - Cost must decrease by updating the model only when and where needed
- To determine the best threshold to detect stable AP
  - The accuracy must increase by using only stable AP for localization

For that

Detect appearance, disappearance, and other environmental change including relocation of AP analyzing available scanned logs

Example of scanned logs

2016-09-03 02:10:3616		
BSSID	ESSID	RSSI
00:09:b4:70:1d:c7	0002softbank	-55
12:09:b4:70:15:f6	FREE_Wi-Fi_PASSPORT	-80
b4:c7:99:16:07:34	Secured_Wi-Fi	-40

Approach

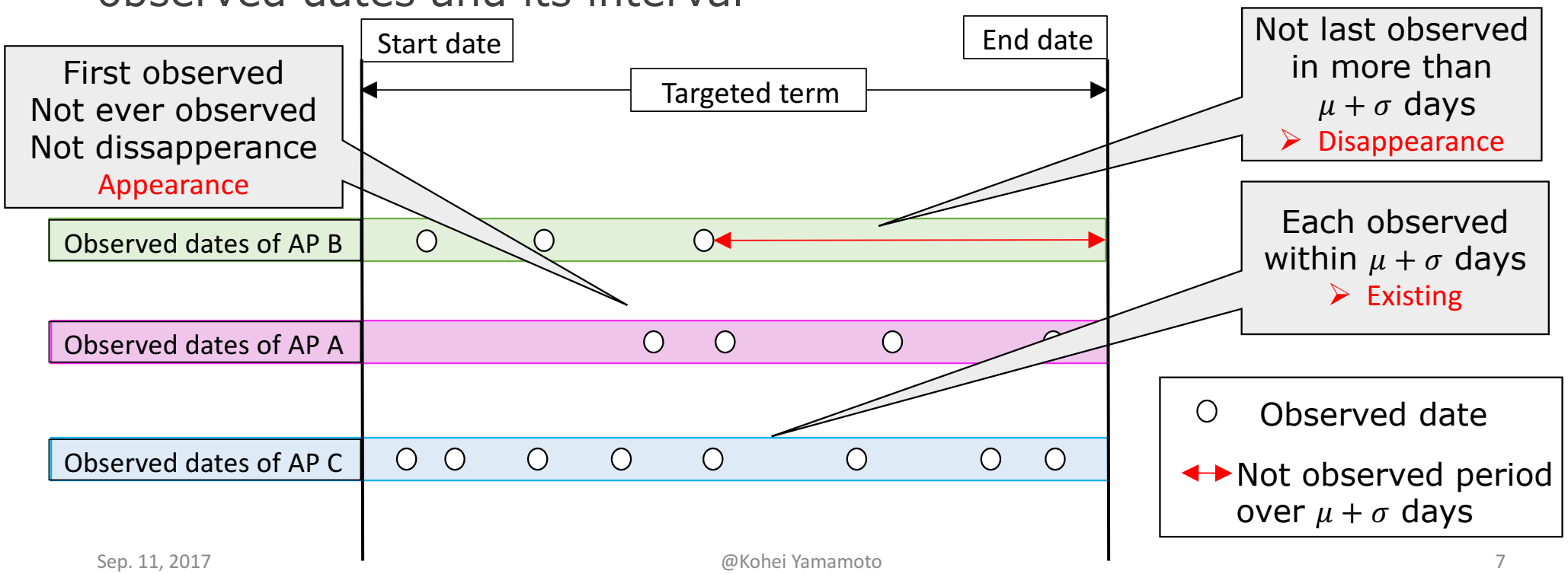
- Detecting appearance and disappearance of AP
  - By chronologically analyzing observed dates and its interval for each AP
- Detecting other environmental change
  - By analyzing changes of co-occurrence of observed AP
- Identifying AP more precisely
  - Filter the application logs by local address and vendor&model code

Approach

- Detecting appearance and disappearance of AP
  - By chronologically analyzing observed dates and its interval for each AP
- Detecting other environmental change
  - By analyzing changes of co-occurrence of observed AP
- Identifying AP more precisely
  - Filter the application logs by local address and vendor&model code

Appearance and disappearance of AP

Determine average (μ) and standard deviation (σ) regarding observed dates and its interval



Approach

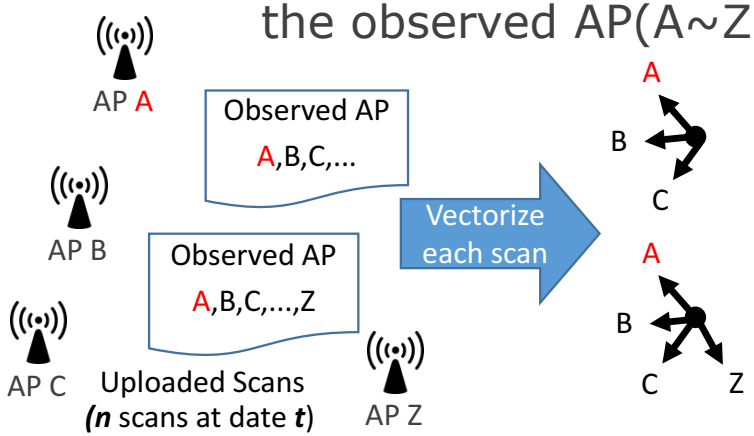
- Detecting appearance and disappearance of AP
  - ➢ By chronologically analyzing observed dates and its interval for each AP
- Detecting other environmental change
  - ➢ By analyzing changes of co-occurrence of observed AP
- Identifying AP more precisely
  - Filter the application logs by local address and vendor&model code

Other environmental change (1/2)

- Co-occurrence will change if AP is relocated or environmental change like obstacle construction happens

➢Express every AP by vector sets

- Generate ***n*** vector sets by vectorizing at the date ***t***
  - Each of n scans observes A
- Each dimension ***C*** consists of vector set expresses the combination with all the observed AP(A~Z) and has binary value



$$\overrightarrow{A_{t,1}} = \left(C_{1_{AA}}, C_{1_{AB}}, C_{1_{AC}}, \cdots, C_{1_{AZ}}\right)$$

Has 0 if AP A is not co-observed with AP Z

$$\vdots$$

$$\overrightarrow{A_{t,n}} = \left(C_{n_{AA}}, C_{n_{AB}}, C_{n_{AC}}, \cdots, C_{n_{AZ}}\right)$$

Has 1 if AP A is co-observed with AP Z

Co-occurrence of observed AP：stable AP tends to be observed with specific AP every time

Other environmental change (2/2)

- Investigate if AP would be observed with specific AP by calculating variance of vector set of AP A($\overrightarrow{\mathrm{A}_{ti}}, \cdots, \overrightarrow{\mathrm{A}_{tn}}$)
- $\overrightarrow{\mathrm{A}_{ti}}$: Vector set of AP A whose date is date ***t***, the number of times observed ***i***
- $\overrightarrow{g_A}$: Center of gravity calculated from vector set of AP A($\overrightarrow{\mathrm{A}_{ti}}, \cdots, \overrightarrow{\mathrm{A}_{tn}}$)

$$\sigma_{\mathrm{A}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left|\overrightarrow{\mathrm{A}_{ti}} - \overrightarrow{g_A}\right|^2$$

Calculate each variance of other AP in the same way

➢AP whose variance is more than certain value could be suspicious

Approach

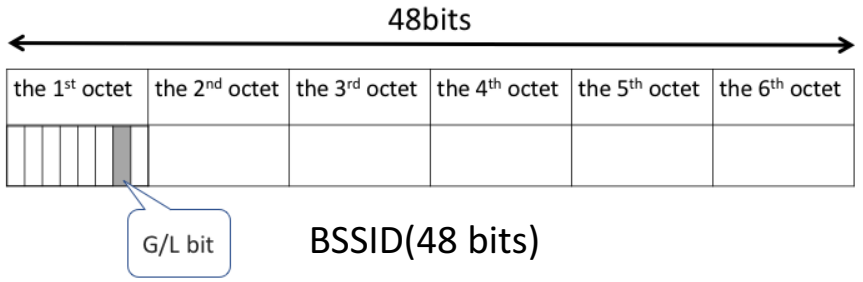
- Detecting appearance and disappearance of AP
  - By chronologically analyzing observed dates and its interval for each AP
- Detecting other environmental change
  - By analyzing changes of co-occurrence of observed AP
- Identifying AP more precisely
  - Filter the application logs by local address and vendor&model code

More precise AP identification (1/2)

- Sometimes hard to identify AP by BSSID
  1. Randomized BSSID
  2. Multiple different BSSID coming from single AP

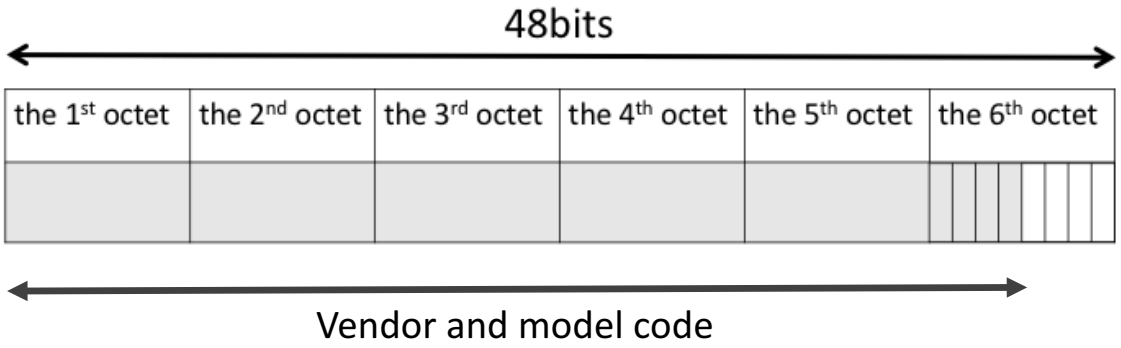
1. Randomized BSSID
   - There can be different AP even if they have the same BSSID
     - BSSID are randomized to protect privacy
   - BSSID whose 2nd bit of 1st octet(called G/L bit) is 1 can be local address

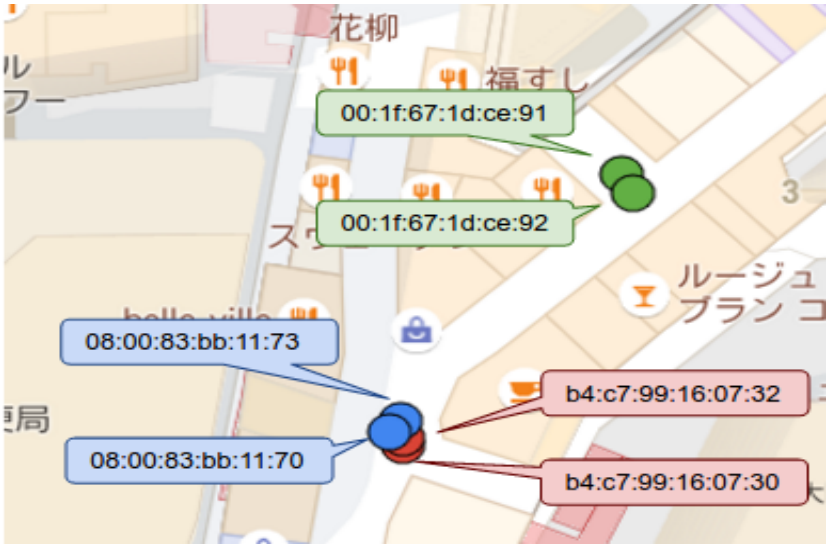
   ➢ Eliminate such BSSID
     - Such BSSID cannot be unique as identifier



More precise AP identification (2/2)

2. Multiple different BSSID coming from single AP
   - If different BSSID show the same vendor and model code, they can be the same AP
   - The sequence of bits from the 1st to 5.5th octet is vendor and model code





Evaluation

- Evaluation environment
  - Analyzed the logs automatically collected from the navigating application "Umechika-navi"[1]
  - AP observed more than once during the targeted term are investigated
    - ➢ Each AP are supposed to be observed twice at least to adapt proposed methods
- Details of the collected logs
  - Collected from Sep.1,2016 to Jan.1,2017(4 months)
  - Totally 21,086 scans
  - Each scan contains BSSID,ESSID,RSSI, and the time
  - ➢ Where it scanned is not known

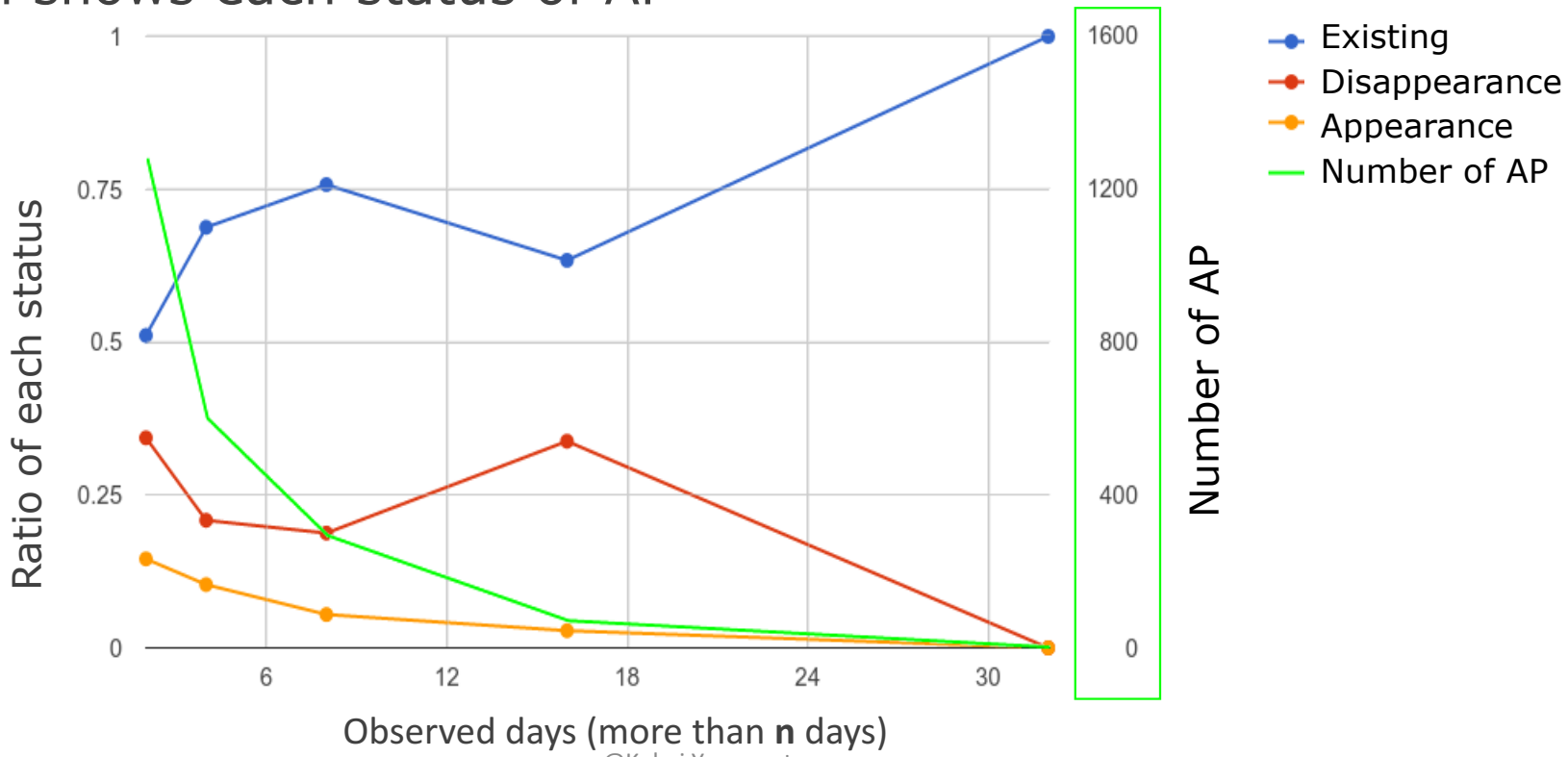
[1] Umechika-navi：navigating application covering an underground district in Umeda, Osaka

Evaluation items

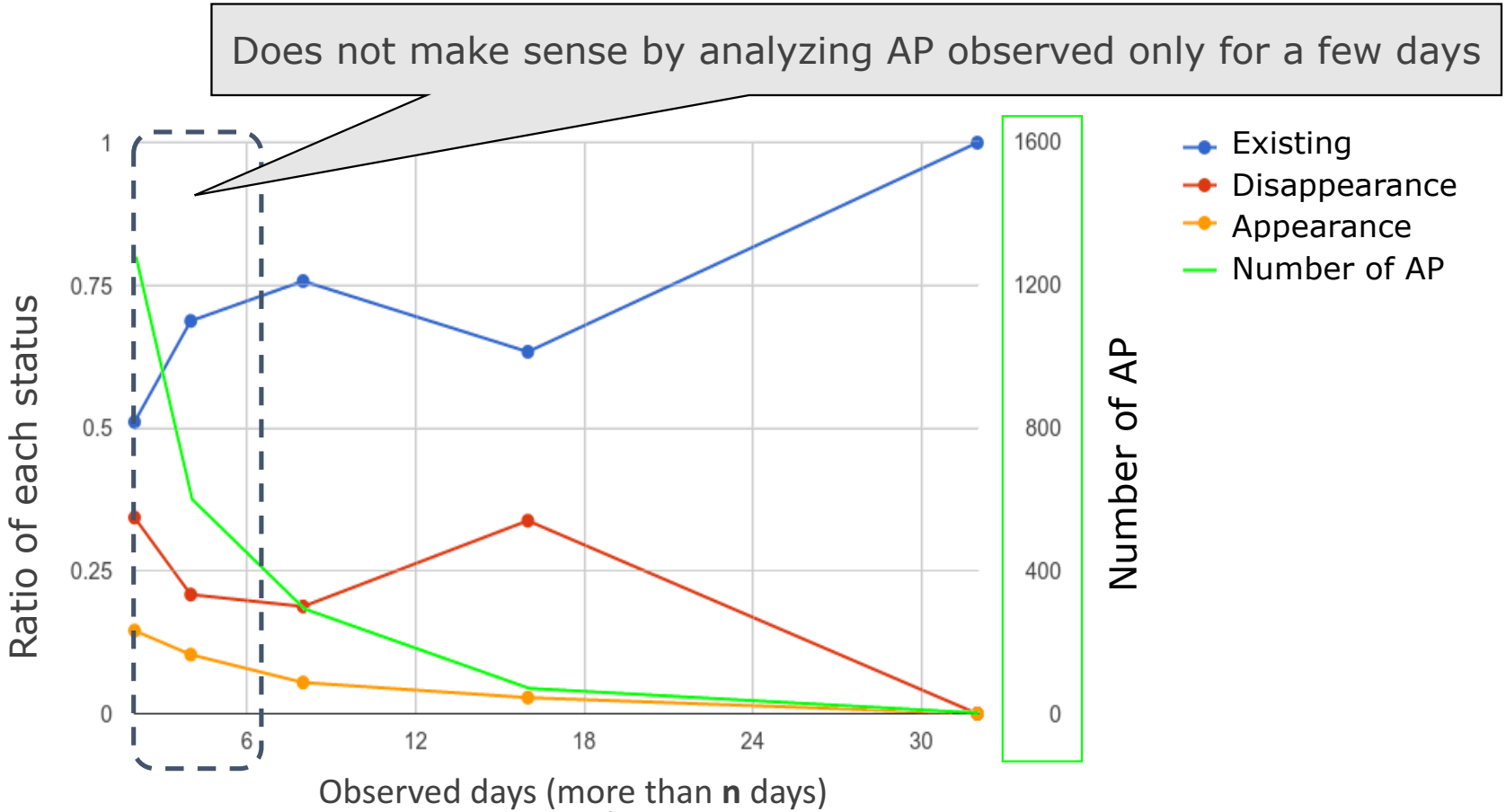
- Detection of appearance and disappearance of AP
  - Classify each AP into appearance, disappearance, and existing by observed days
- Detection of other environmental change
  - Investigate each variance of AP by the number of times observed and days
  - Determine the variance considering fluctuation caused by temporal AP

Evaluation Results (appearance and disappearance)

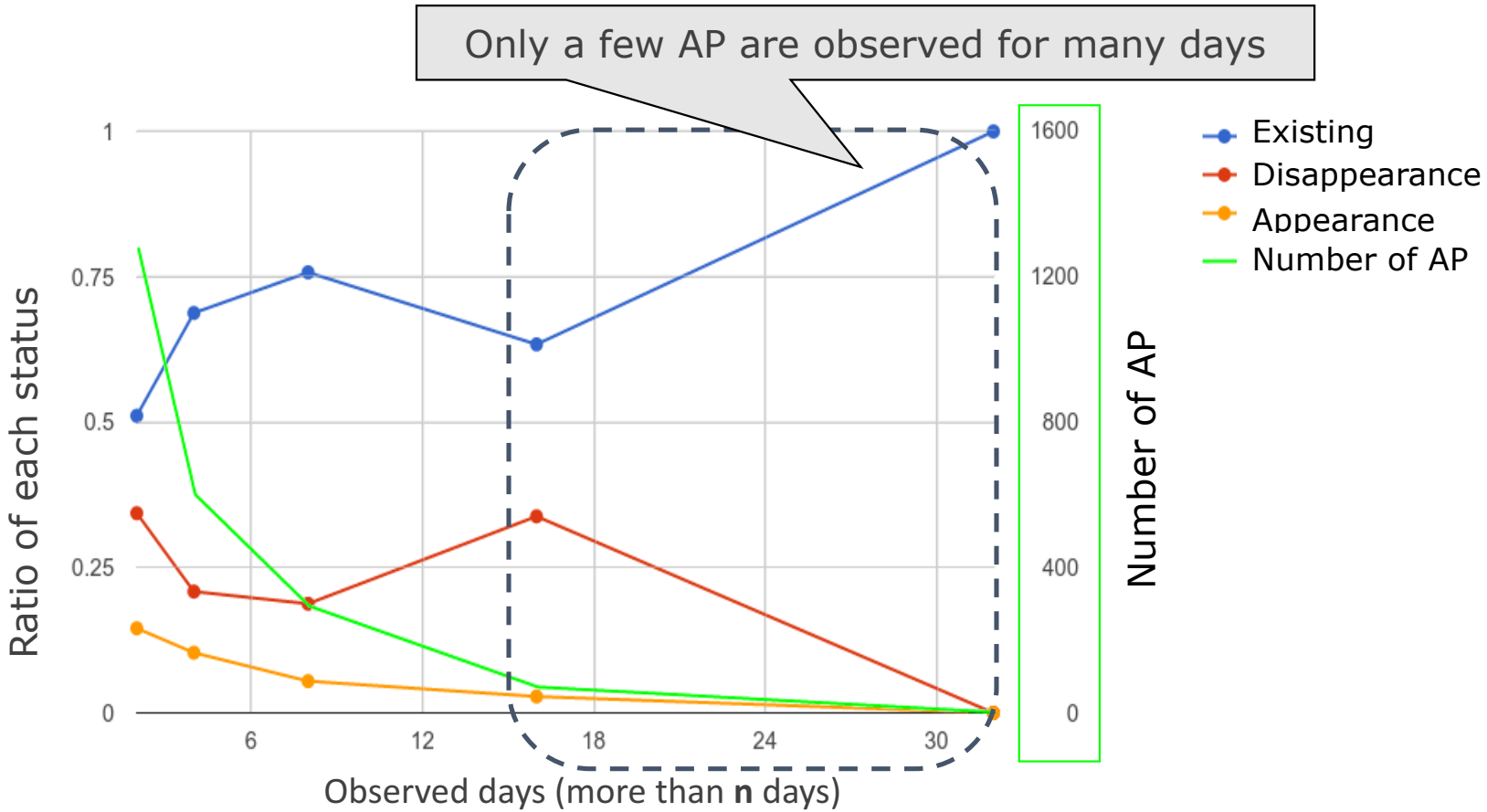
- Graph shows each status of AP



Evaluation Results (appearance and disappearance)

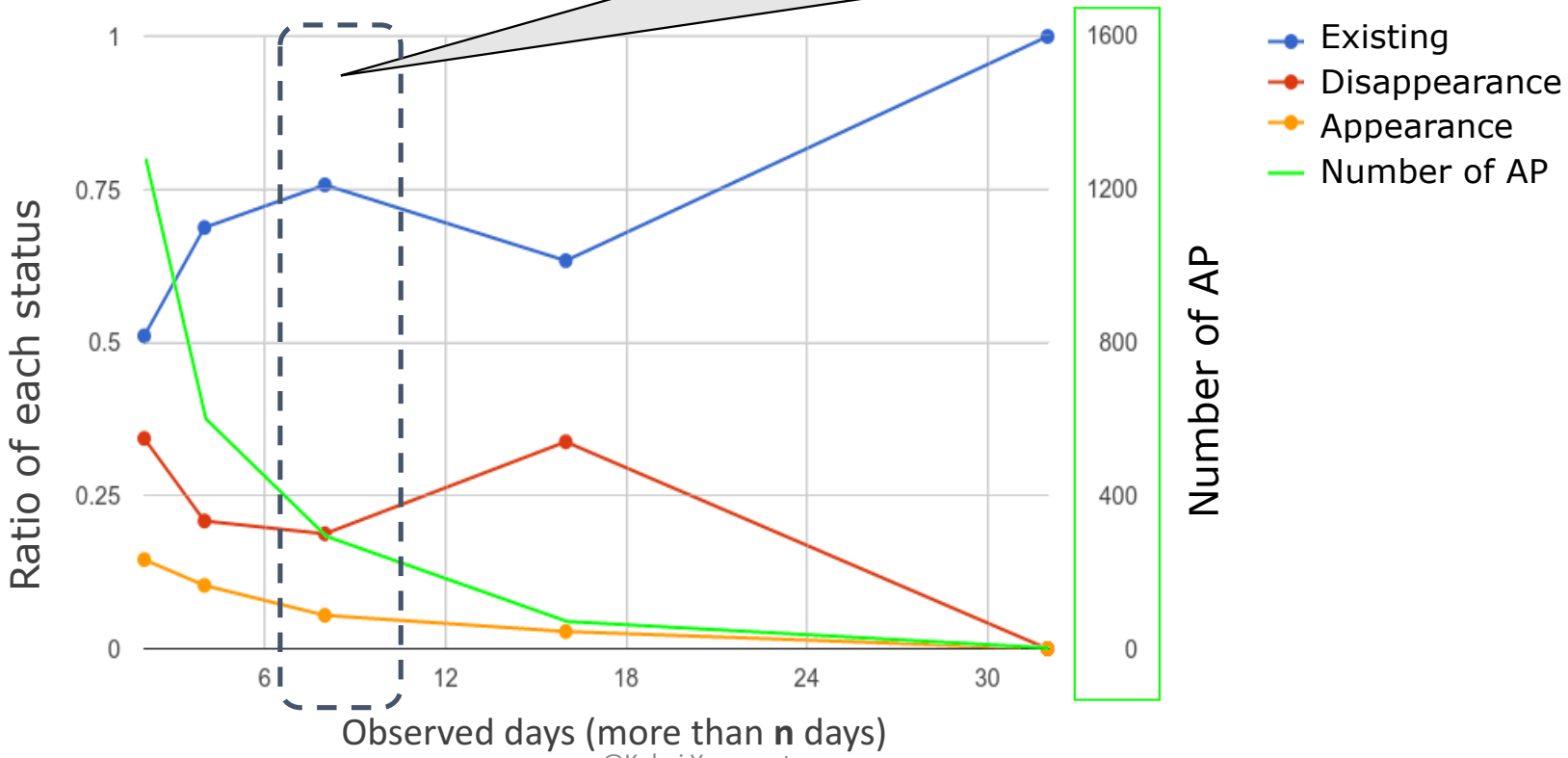


Evaluation Results (appearance and disappearance)



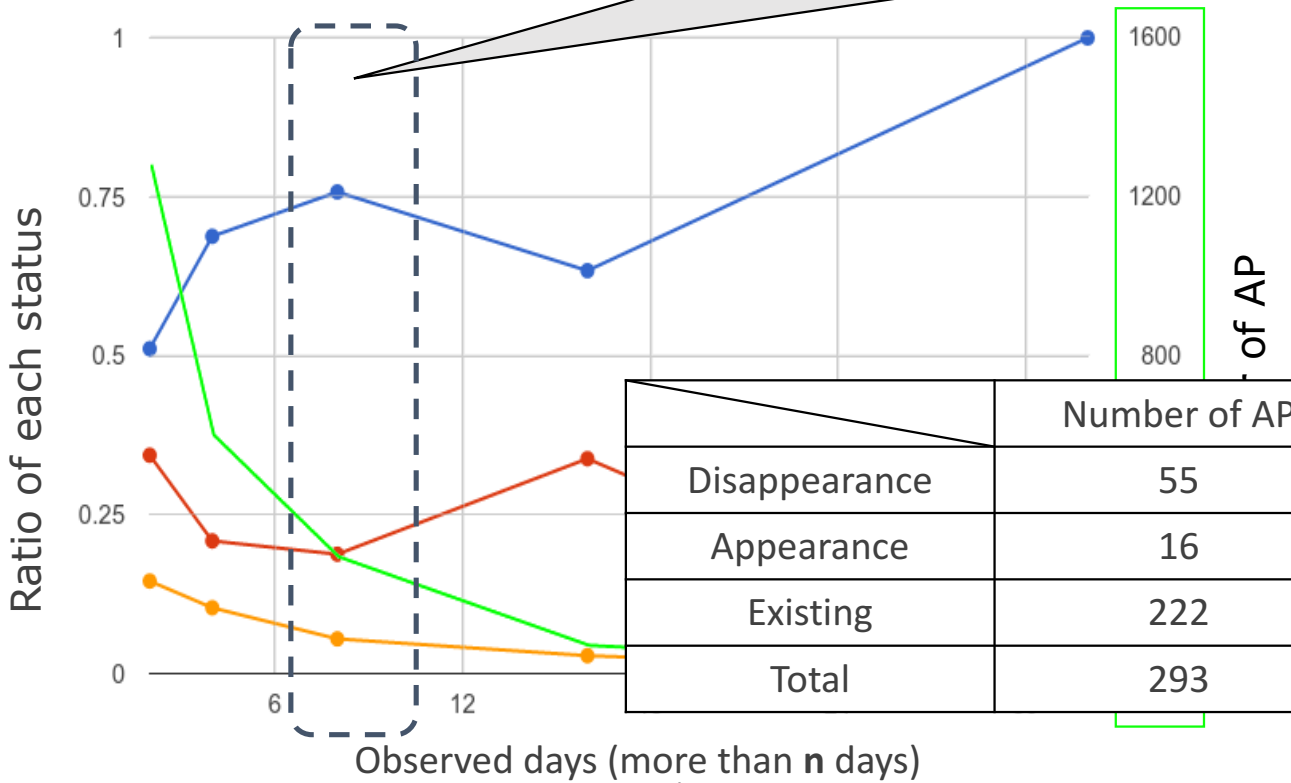
Evaluation Results (appearance and disappearance)

AP observed for more than 8 days are more close to reality



Evaluation Results (appearance and disappearance)

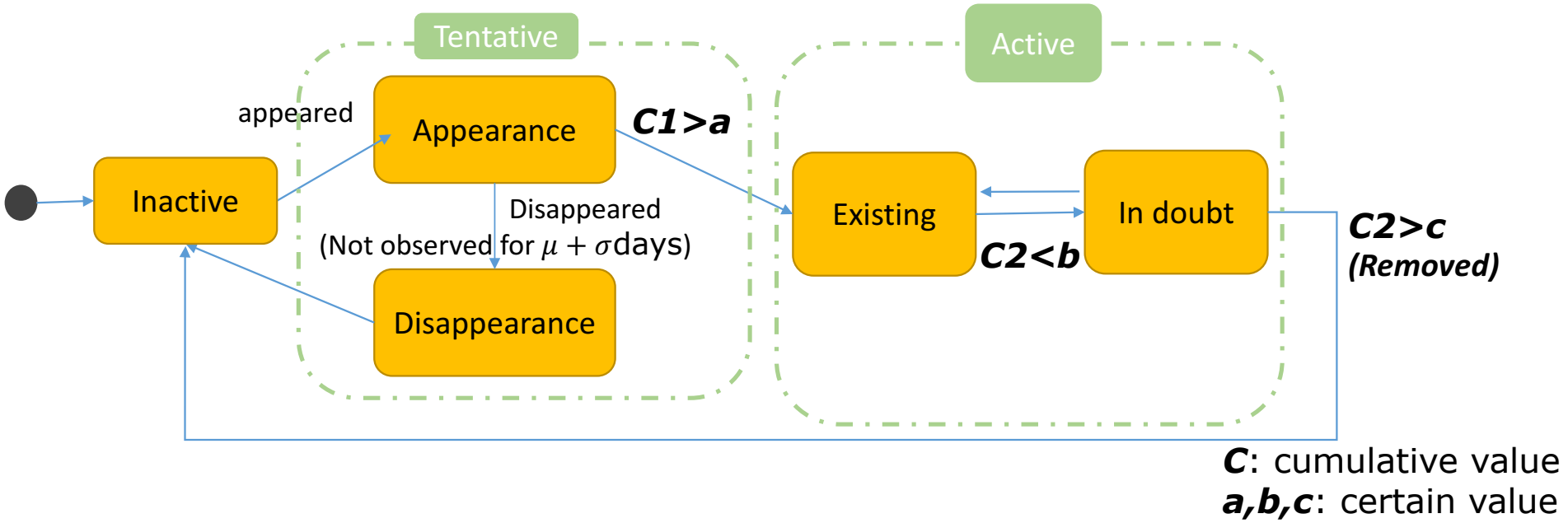
AP observed for more than 8 days are more close to reality



	Number of AP	Ratio of AP
Disappearance	55	19[%]
Appearance	16	5[%]
Existing	222	76[%]
Total	293	100[%]

State Transition Diagram to manage each AP

- Might be better to use STD to manage each AP's status
  - The way we did is a bit too discrete to manage each AP's status
  - With introducing some cumulative values



Evaluation items

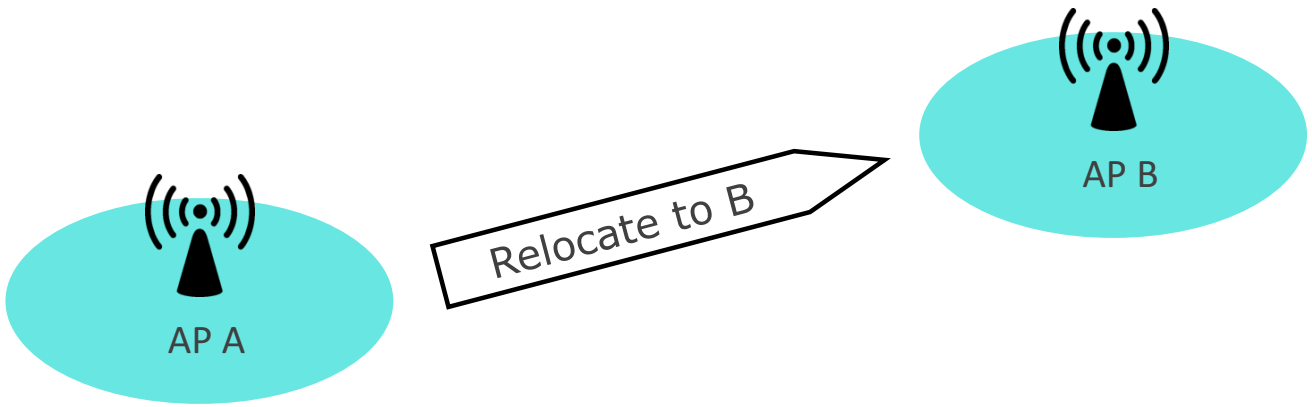
- Detection of appearance and disappearance of AP
  - Classify each AP into appearance, disappearance, and existing by observed days
- Detection of other environmental change
  - Investigate each variance of AP by the number of times observed and days
  - Determine the variance considering fluctuation caused by temporal AP

Variance of AP (1/5)

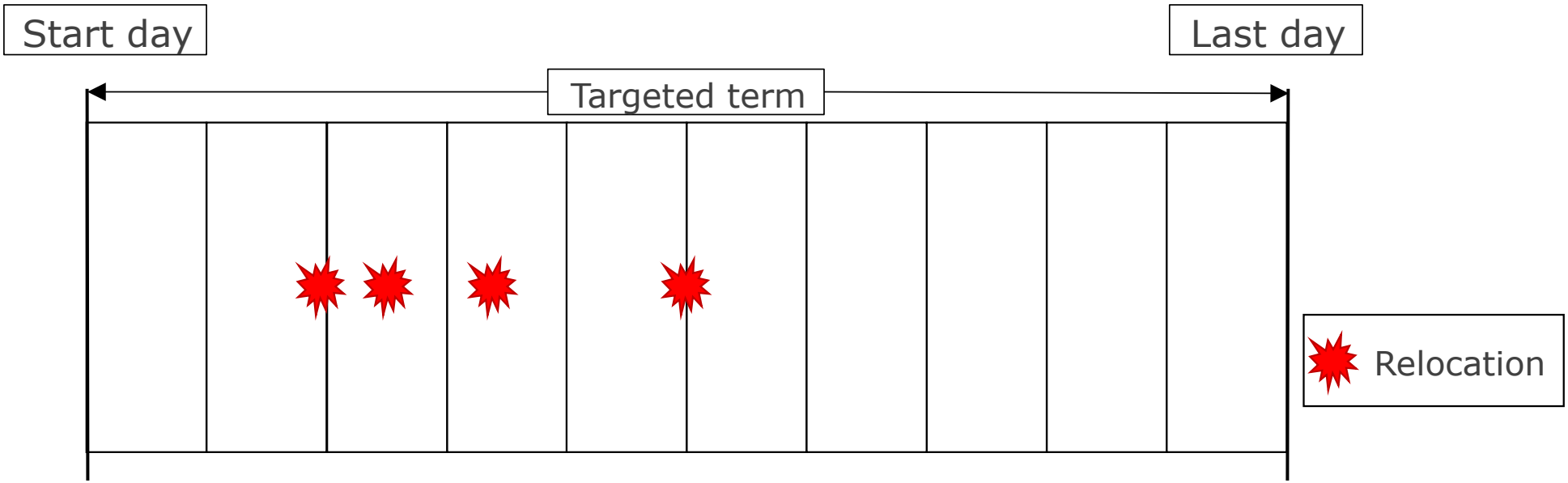
- 11 scans observed on average per day and nine AP are co-occurred through the targeted term

➢Select two AP(A, B) observed stably through the targeted term

➢Calculate variance simulating relocation from A to B

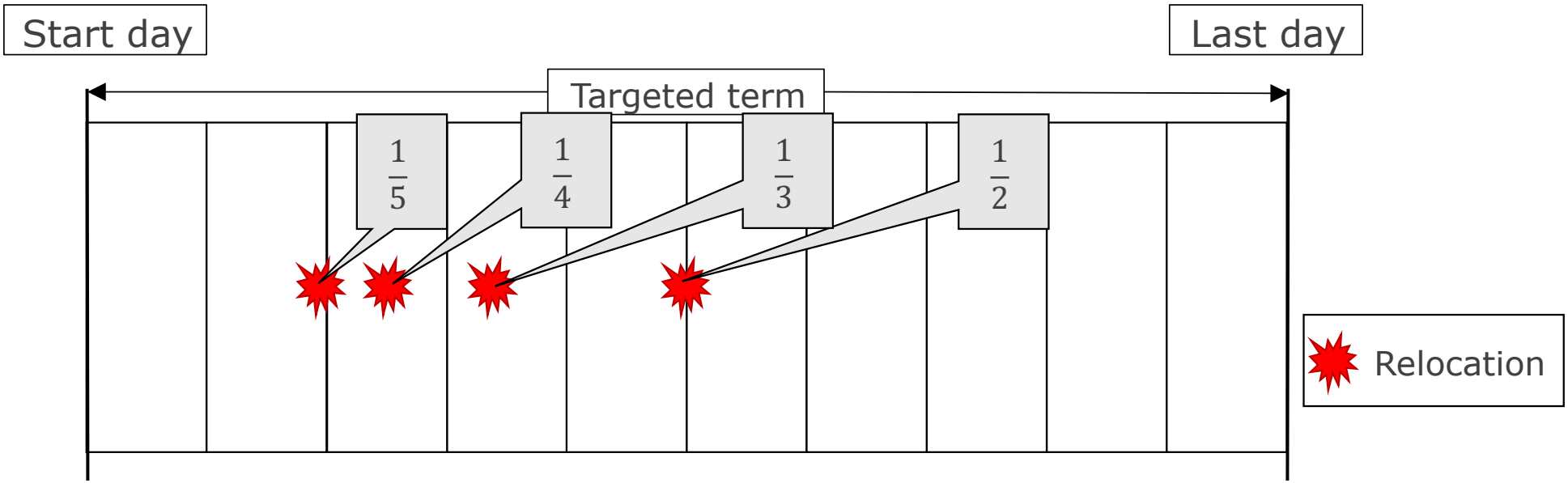


Variance of AP (2/5)



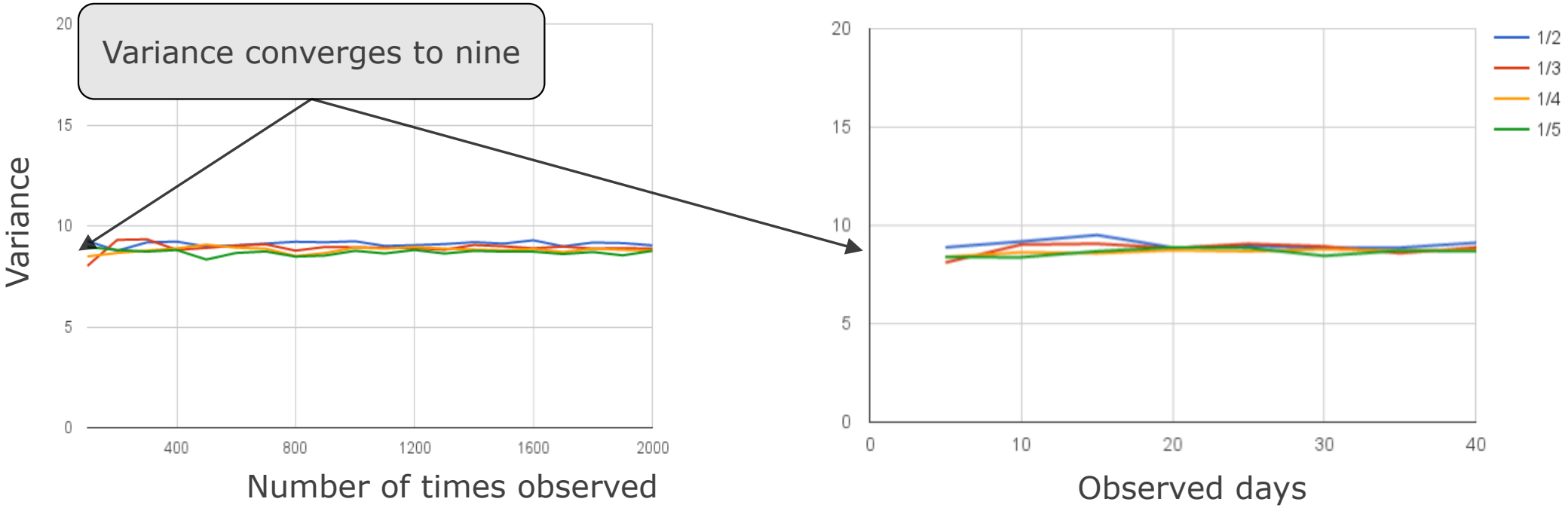
- Variance is plotted if relocation occurs at $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$, and $\frac{1}{5}$ of the targeted term

Variance of AP (2/5)



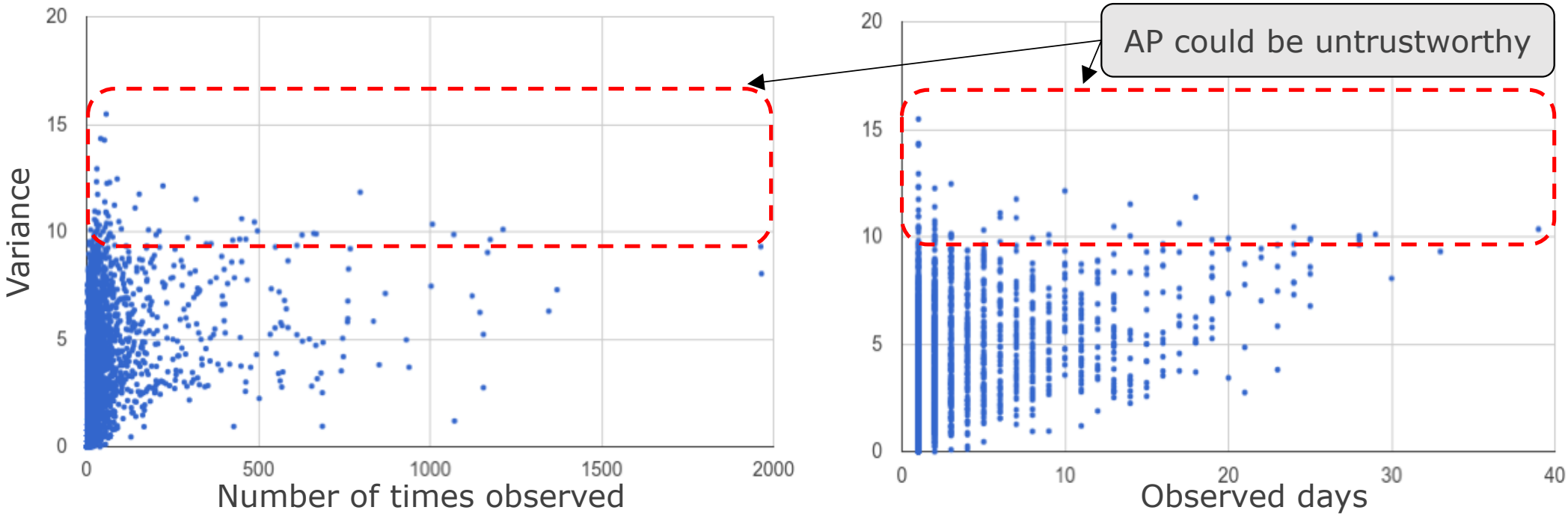
- Variance is plotted if relocation occurs at $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$, and $\frac{1}{5}$ of the targeted term

Variance of AP (3/5)



- Variance converges to nine as times and observed days increase
- AP whose variance is more than nine are suspicious

Variance of AP (4/5)



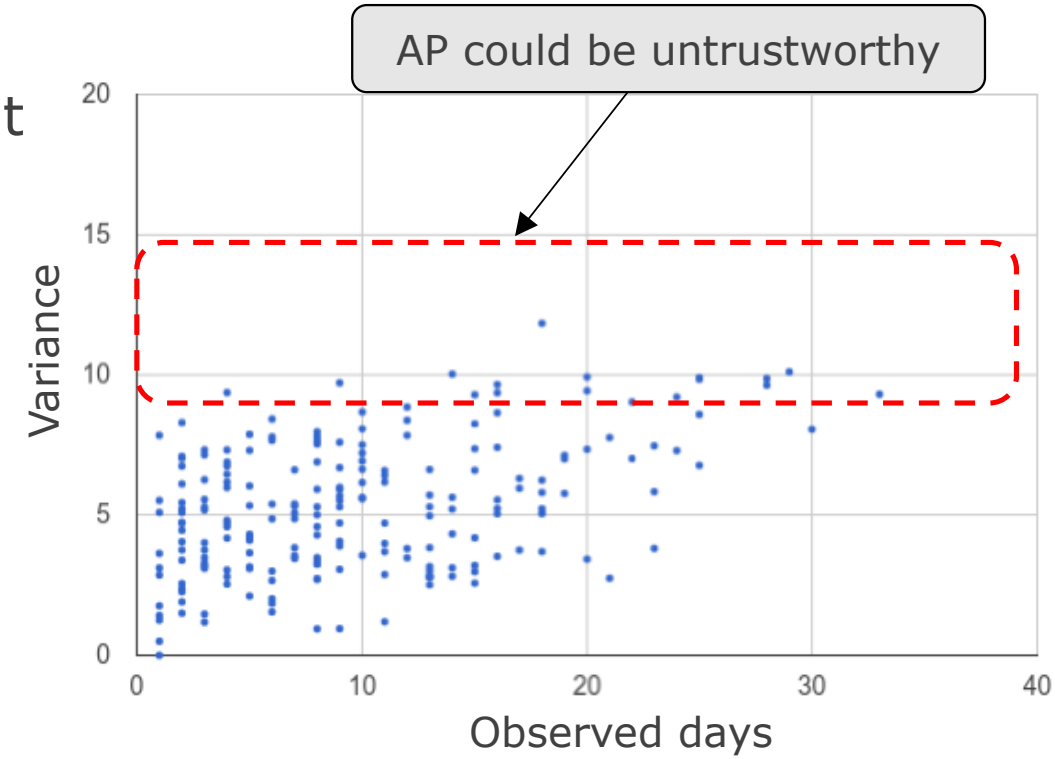
Results show almost same in the both way

Variance of AP(5/5)

- AP installed by carriers thought to be stable

But

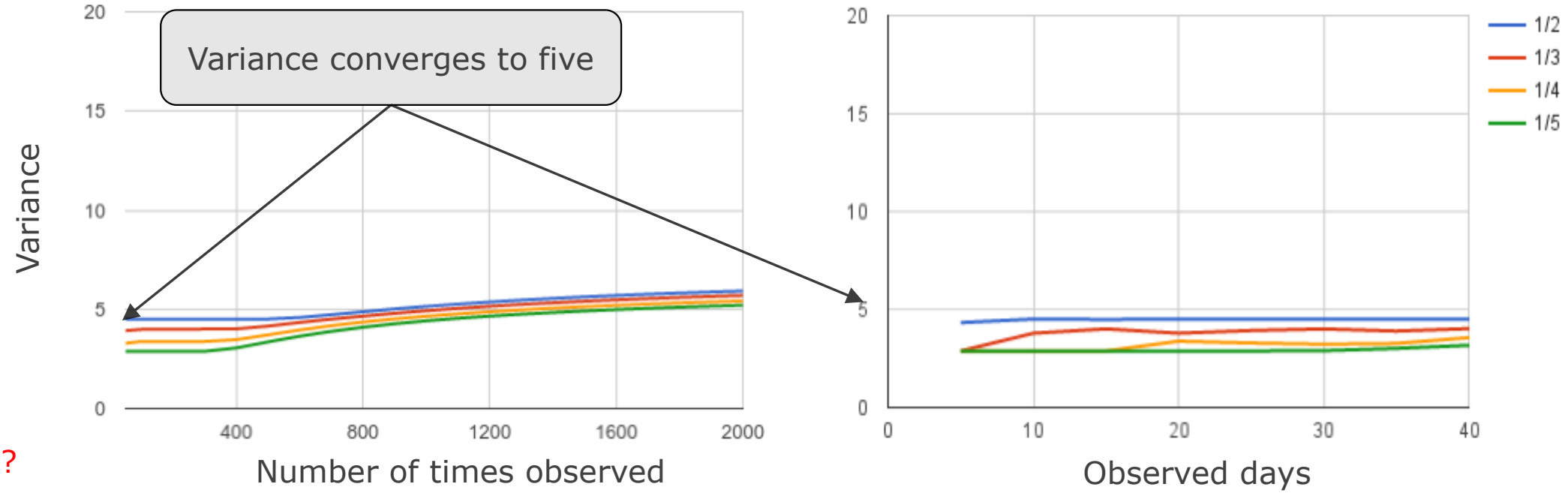
- Not all of the AP are suitable for localization



Summary

- Background
  - Model for localization deteriorates as time passes
  - Cannot evaluate when and to update model
  - Cannot evaluate where and to update model
- Approach
  - Detecting appearance and disappearance of AP by determining average and standard deviation of observed dates and its interval
  - Detecting relocation of AP by analyzing changes of co-occurrence of observed AP
- Evaluation
  - Showed possibility of Detecting AP not suitable for localization only by analyzing collected logs
  - Showed possibility of Detecting 16 possibly appeared AP and 55 disappeared AP during four months
  - AP whose variance shows more than nine could be suspicious
  - Even detected AP installed by carriers
- Future work
  - Detect AP which works on and off
  - Update localization model using collected logs

Variance of AP [Theoretical value]



は？

いるのか

- Variance converges to five as times and observed days increase