Detection of MMORPG Bots Based on Behavior Analysis

Yoshitaka Kashifuji

Ruck Thawonmas

Intelligent Computer Entertainment Lab, ISE, Ritsumeikan University

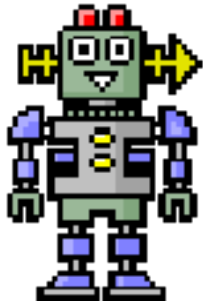
Kuan-Ta Chen

Multimedia Networking and Systems Lab, IIS, Academia Sinica

ACE 2008, 4 DEC 2008

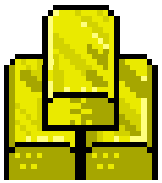
Background & Objectives

- Online game market size ↑
- Misconduct ↑
  - BOT
  - RMT
- Damage of using unauthorized tools
  - Game balance collapses
  - Play motivation ↓
  - Distrust of publishers ↑



Item collected by BOT

Cashes earned with RMT strengthens characters



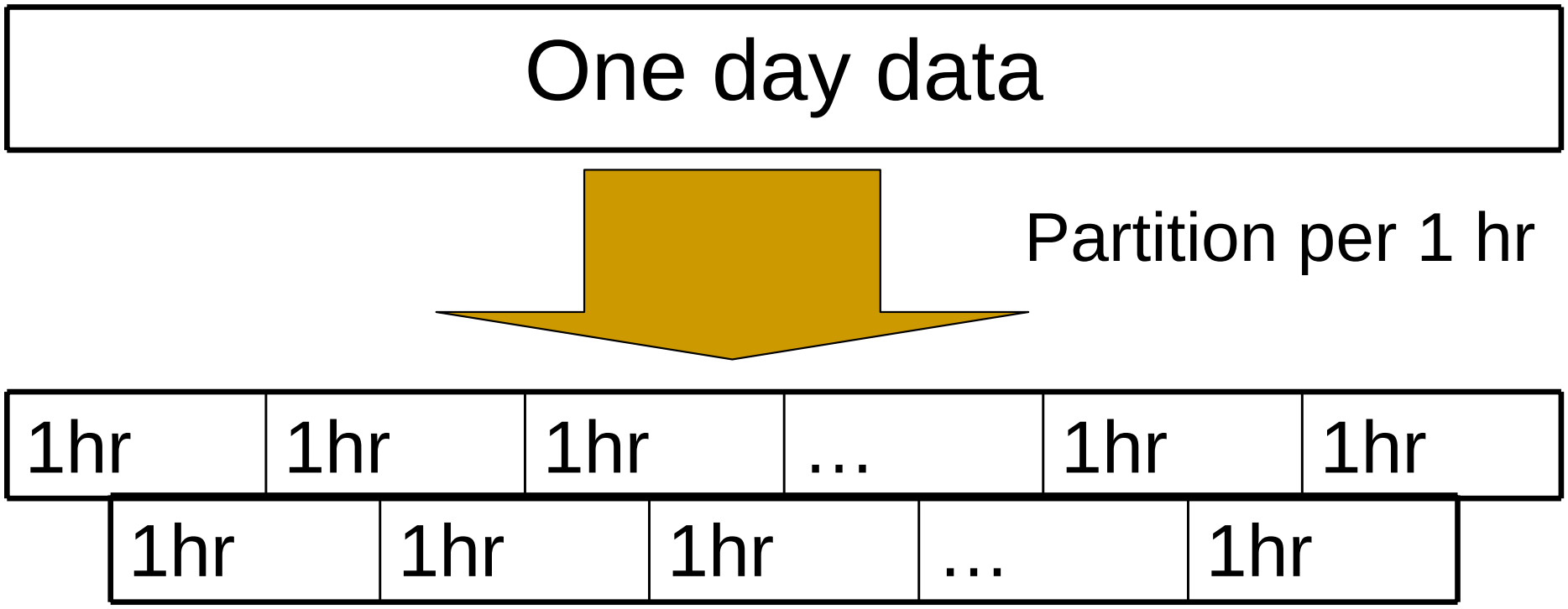
Data Set

- Amount
  - Use of unauthorized macro tool

    BOT 7 characters
  - **Human** 7 characters
  - Total 14
- Log acquisition time:
  - Three days in avg
- Character level:
  - Approximately same

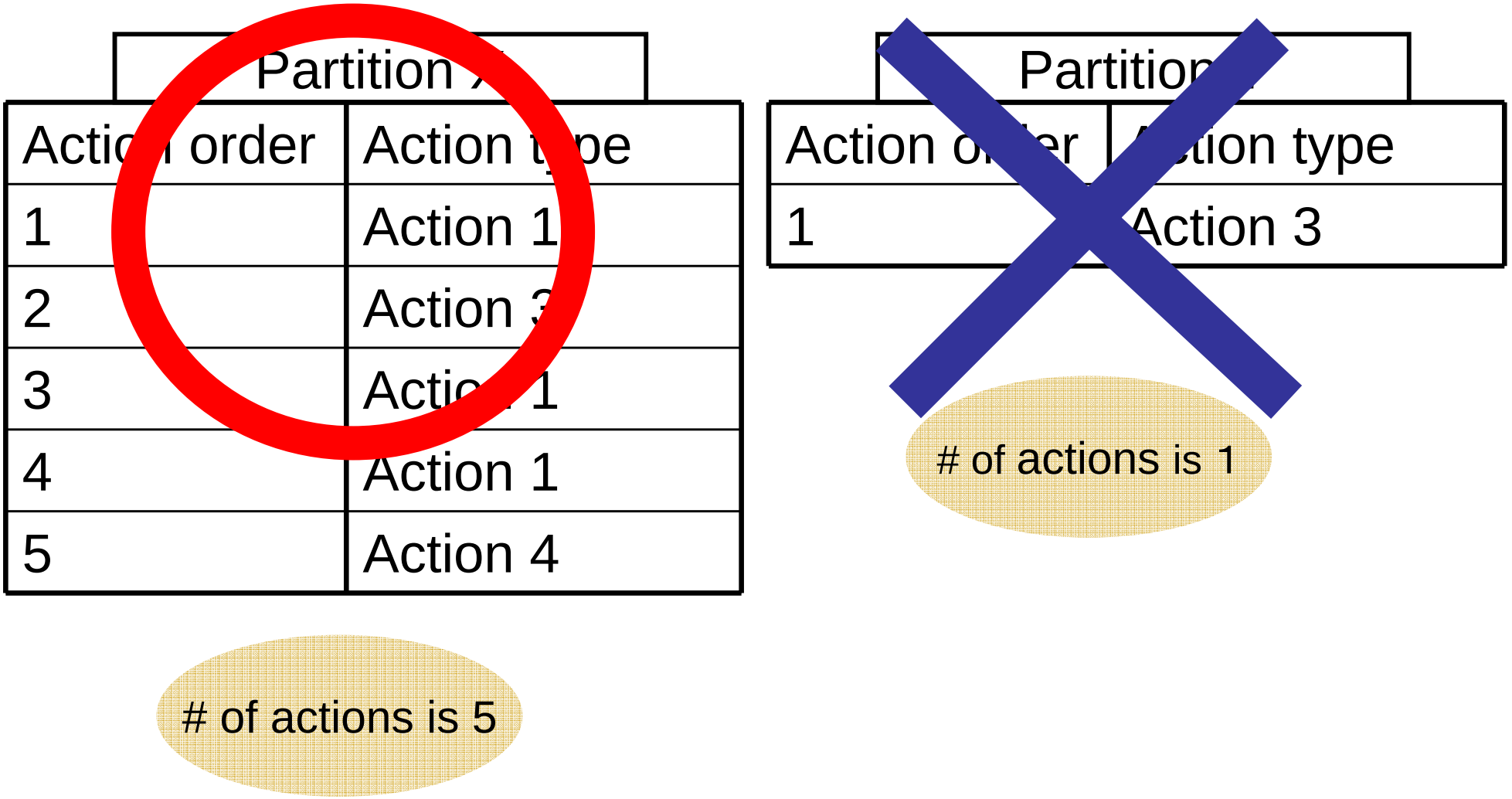
CABAL ONLINE
provided by Gamepot Inc

Data partition(overlap)



Data filtering

- Delete partitions with # of actions below 1

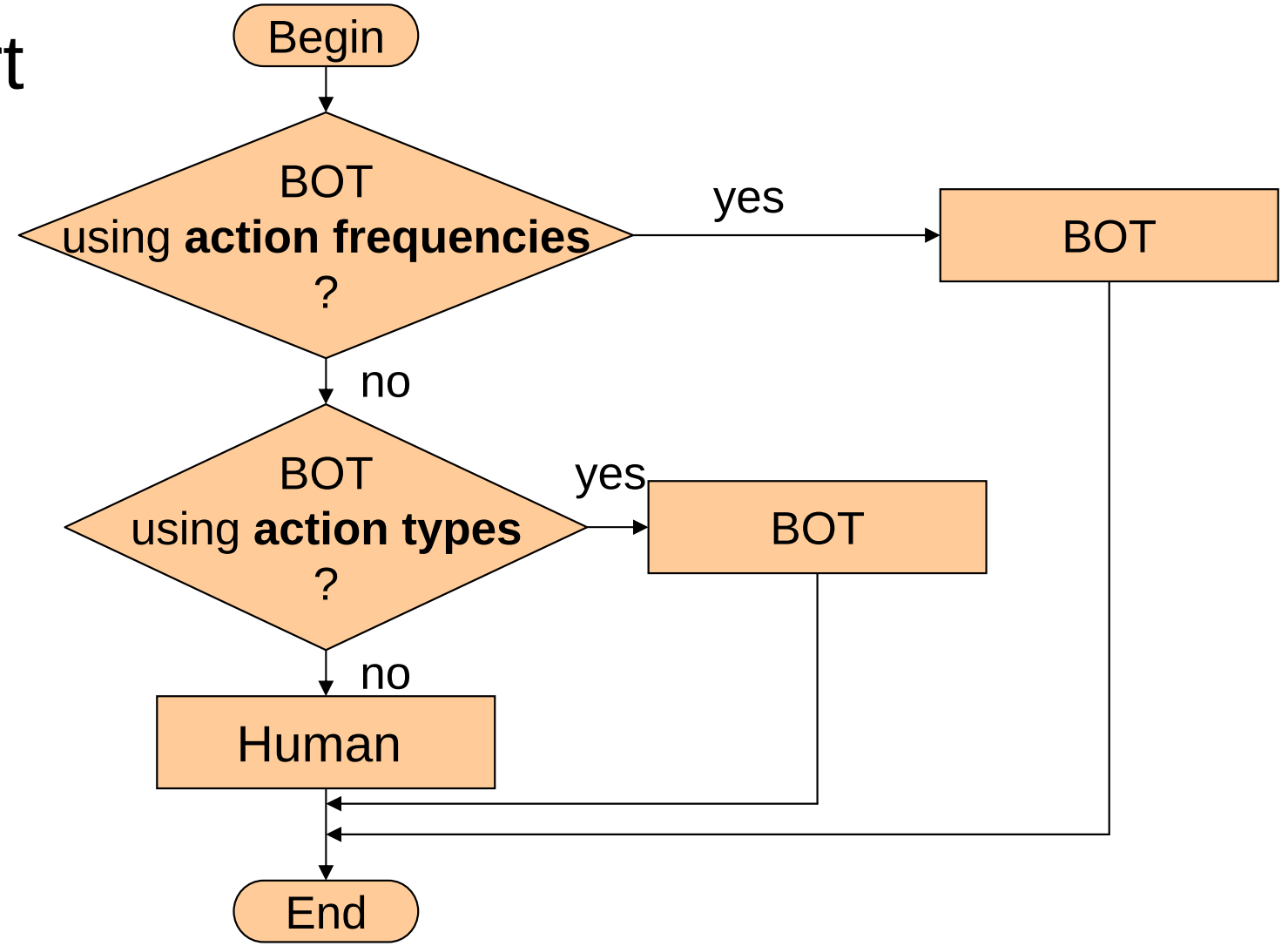


Evaluation method

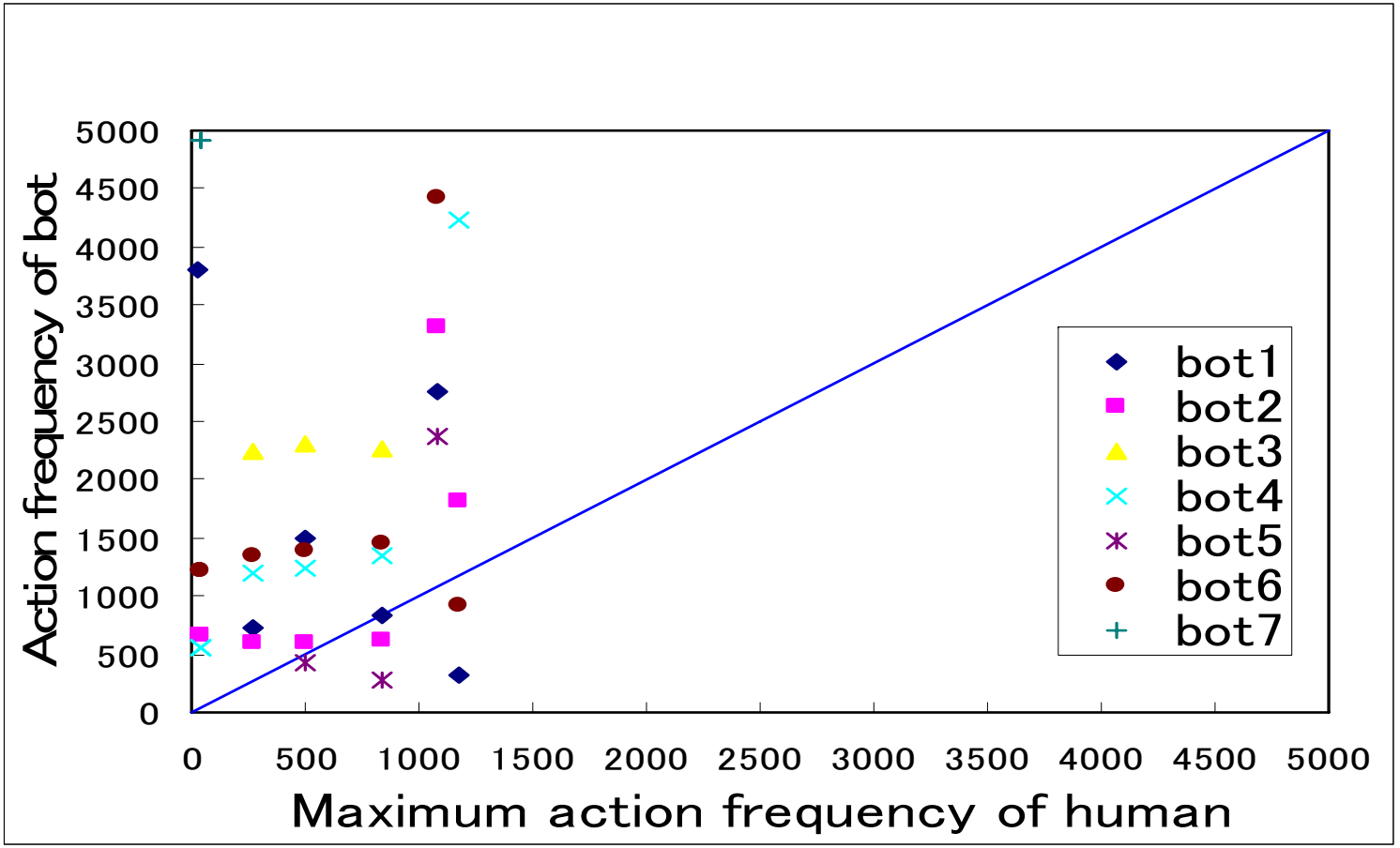
- Training data　12 characters
  - BOT　6
  - Human 6
- Test data（unknown data）　2 characters
  - BOT　1
  - Human　1
- Cross-validation
  - 49 combinations

Classification using action frequencies and types

- Flowchart



Scatter plot of action frequencies of bots and maximum action frequencies of humans



- X axis : Max frequency of each action by Human
- Y axis : Frequency of each selected action of each BOT
- Selected actions : Actions whose frequencies are above average for the corresponding BOT

Classification with action frequencies

- *Frequencies of selected actions for a BOT are larger than those of Humans*
- Training method (to decide parameters)
  1. Obtain max frequency of each action "a", "**max_human_freq(a)"**, from all Humans in training data
  2. Obtain threshold ρ from all BOTs in training data
- Classify unknown player
  1. **Actions with frequencies above average for the unknown player**
  2. **Actions with max_human_freq(a) != 0**
  3. **Actions with frequencies above max_human_freq(a)*** ρ

If at least one action exists satisfying all conditions →BOT,
else → Pending

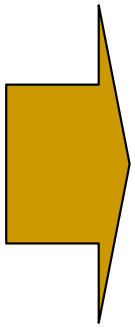
- Pending player will be classified later using action types

Ex : Training method using action frequencies

Obtain "max_human_freq(a)"

Training Data Human

player \ Action	A	B	C
Human1	3	20	0
Human2	2	30	0
Human3	0	0	0



max_human_freq

Action	A	B	C
#	3	30	0

Ex: Training method using action frequencies

Obtain ρ

1. Select freq-above-avg actions for each training BOT
2. Exclude actions with max_human_freq = 0
3. Further select actions for each training BOT with BOT frequency / max_human_freq ≧ 1
4. List up the minimum ratio from the selected actions, if any, for each BOT
5. Use the median from the listed ratios in 4

Training data BOT

player \ Action	A	B	C	avg
BOT1	~~2~~	1000	~~0~~	334
BOT2	~~0~~	10	~~0~~	5
BOT3	31	40	~~20~~	30
BOT4	3	~~1~~	~~2~~	2

max_human_freq

Action	A	B	C
#	3	30	0

Ex: Training method using action frequencies

Obtain ρ

1. Select freq-above-avg actions for each training BOT
2. Exclude actions with max_human_freq = 0
3. Further select actions for each training BOT with BOT frequency / max_human_freq ≧ 1
4. List up the minimum ratio from the selected actions, if any, for each BOT
5. Use the median from the listed ratios in 4

Training data BOT

Player \ Action	A	B	C
BOT1	×	1000/30=33.3	×
BOT2	×	10/30=0.3	×
BOT3	3/3=1	40/30=1.3	×
BOT4	3/3=1	×	×

max_human_freq

Action	A	B	C
#	3	30	0

List

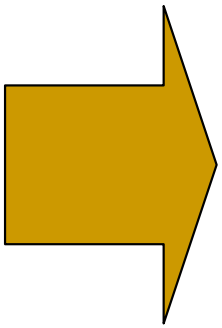
33.3, 1.3, 1

$\rho = 1.3$

Ex: Test using action frequencies

Human Training data

BOT Training data



max_human_freq

Action	A	B	C
#	3	30	0

ρ =1.3

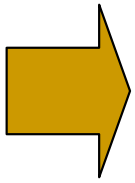
Test data (Unknown data)

Action	A	B	C	Avg
#	~~2~~	100	~~0~~	34

1. Actions with frequencies above average for the unknown player

2. Exclude actions with max_human_freq = 0

3. Actions with frequencies above max_human_freq(a)* ρ

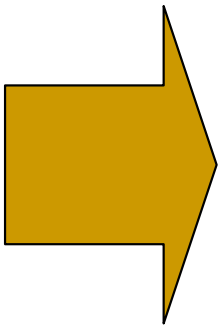


This is BOT

Ex: Test using action frequencies

Human Training data

BOT Training data



max_human_freq

Action	A	B	C
#	3	30	0

ρ =1.3

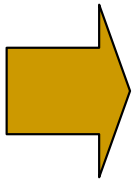
Test data (Unknown data)

Action	A	B	C	Avg
#	1	4	1	3

1. Actions with frequencies above average for the unknown player

2. Exclude actions with max_human_freq = 0

3. Actions with frequencies above max_human_freq(a)* ρ

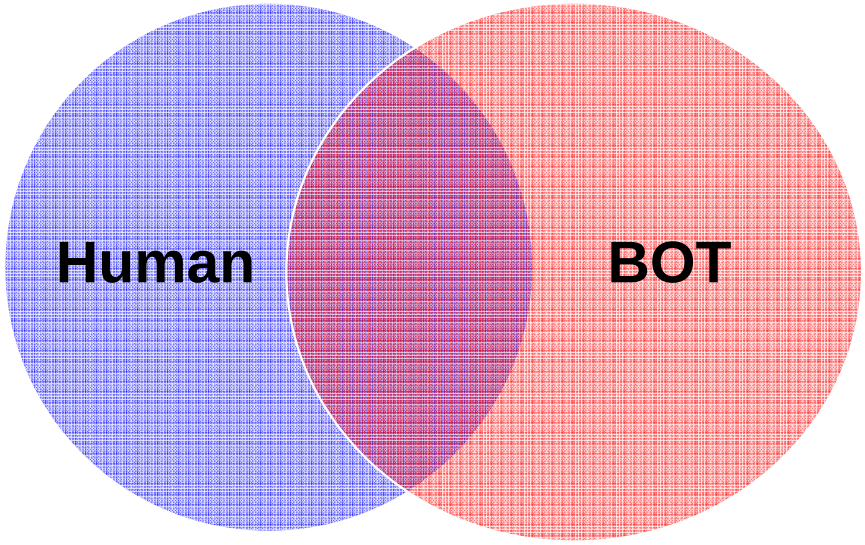


This is Pending

of action types of BOT and Human

- The mean and standard deviation of action type # of BOTs and Humans

Character Type	MEAN	STD
BOT	32.71	8.99
Human	54.28	13.52

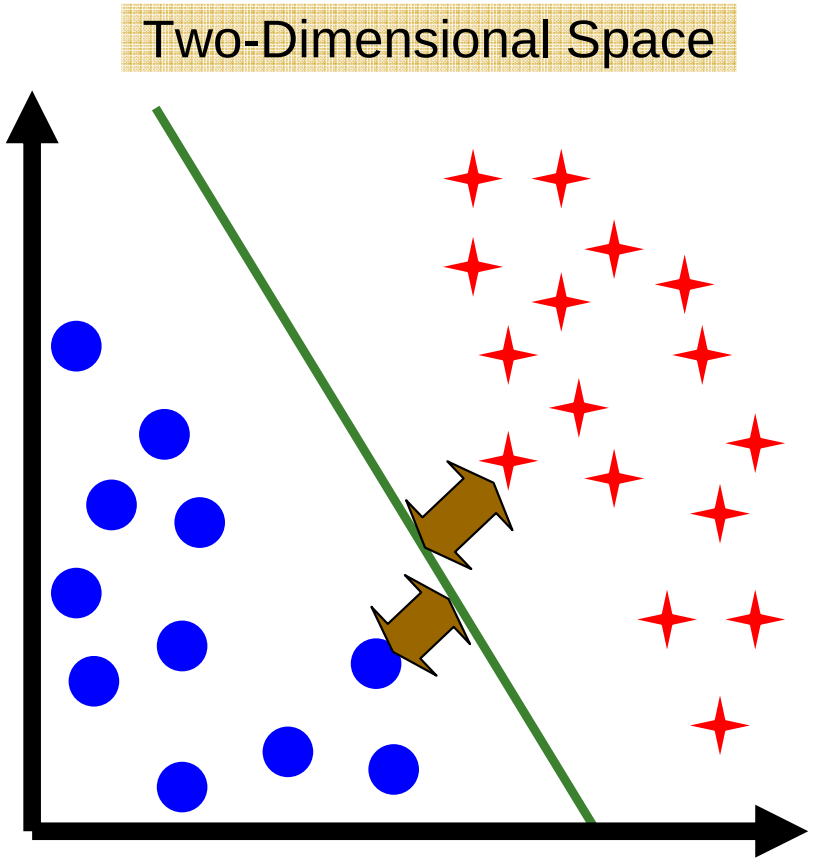


- BOT executes specific actions

→Doing similar sets of actions

Training method using action type

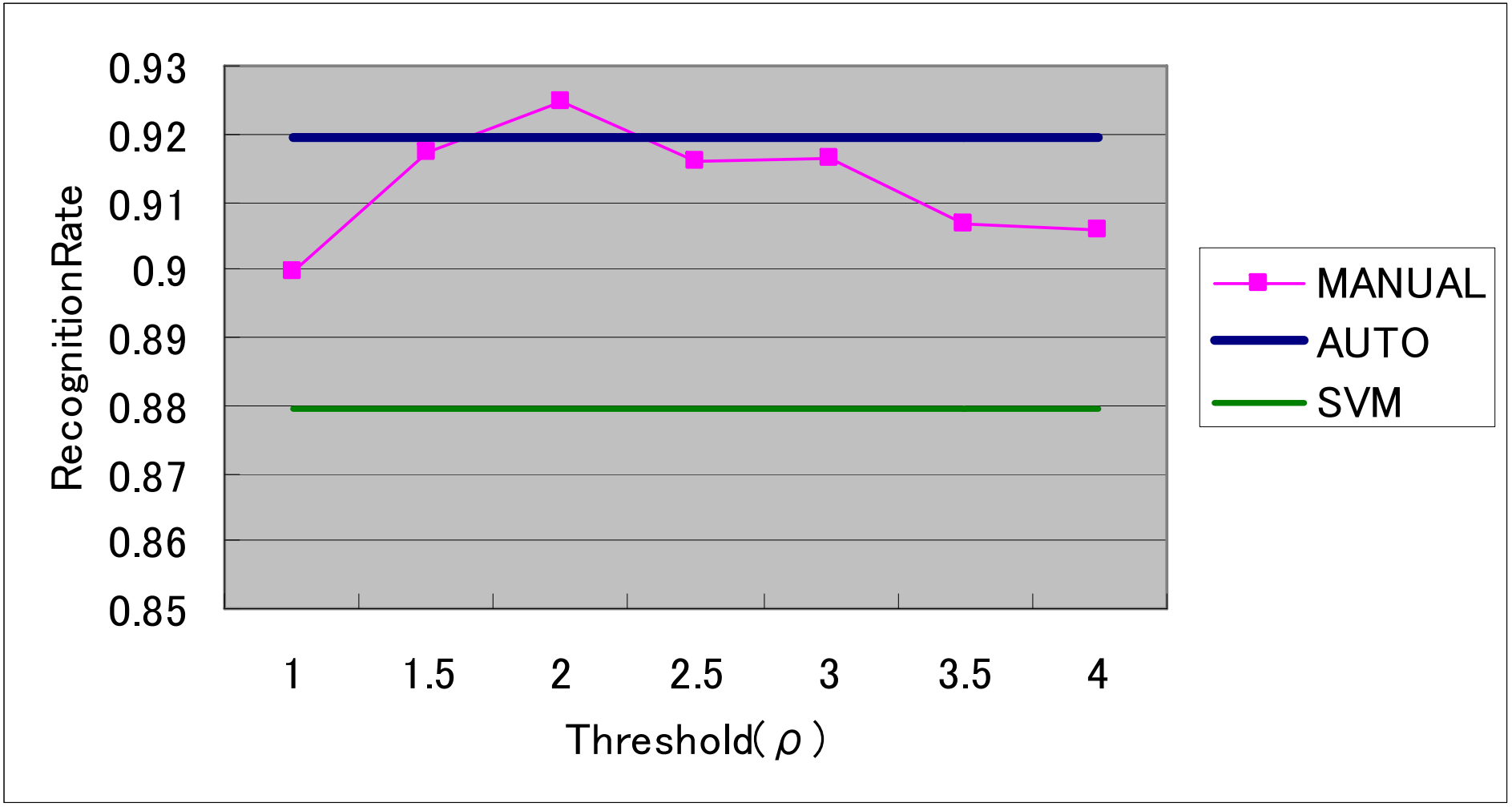
- Support Vector Machine (SVM)
  - Linear kernel.



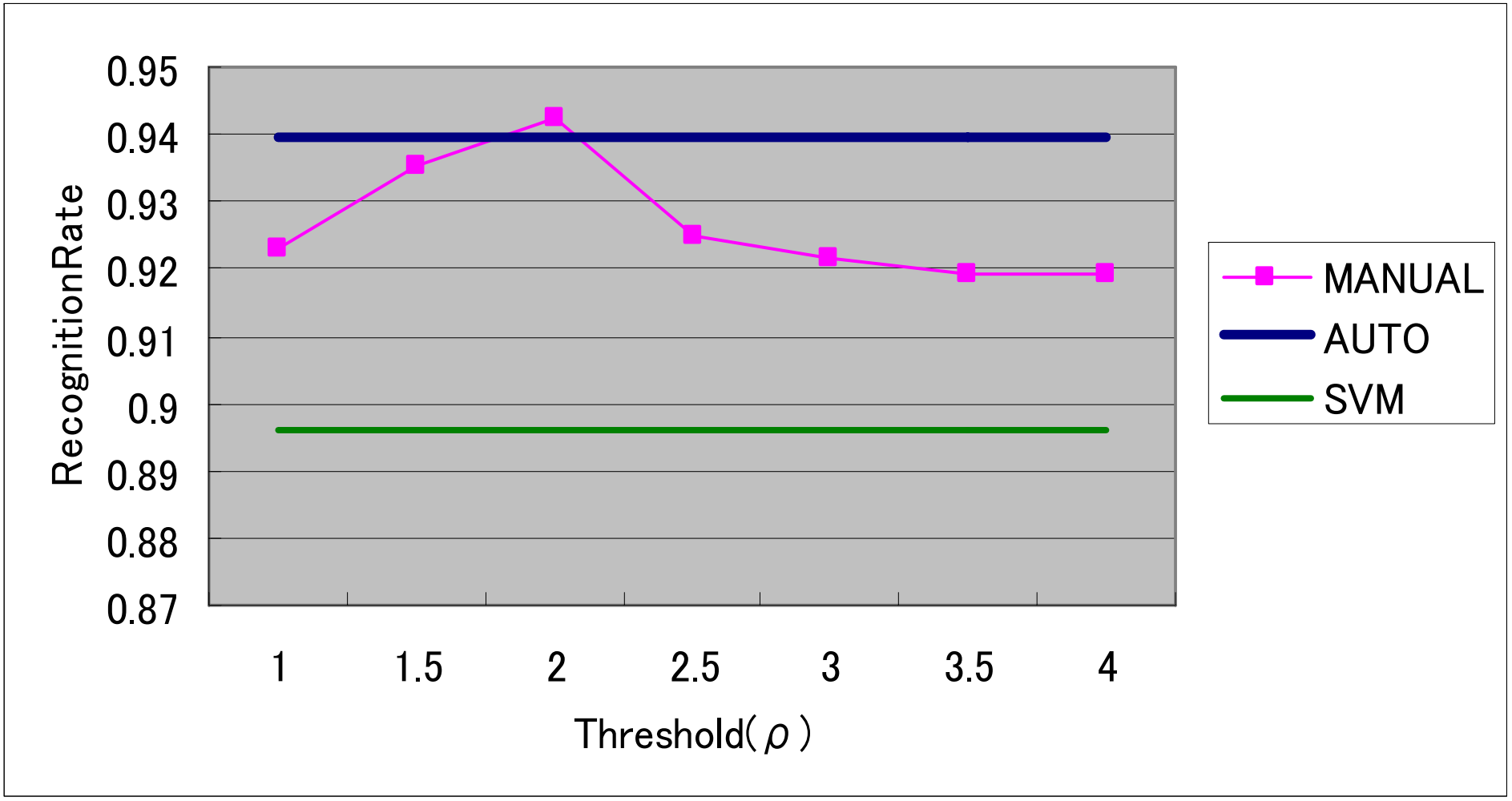
Action order	Action type
1	Action A
2	Action C
3	Action A
4	Action R
5	Action C

Action	A	B	C	…	Z
Training data	1	0	1	…	0

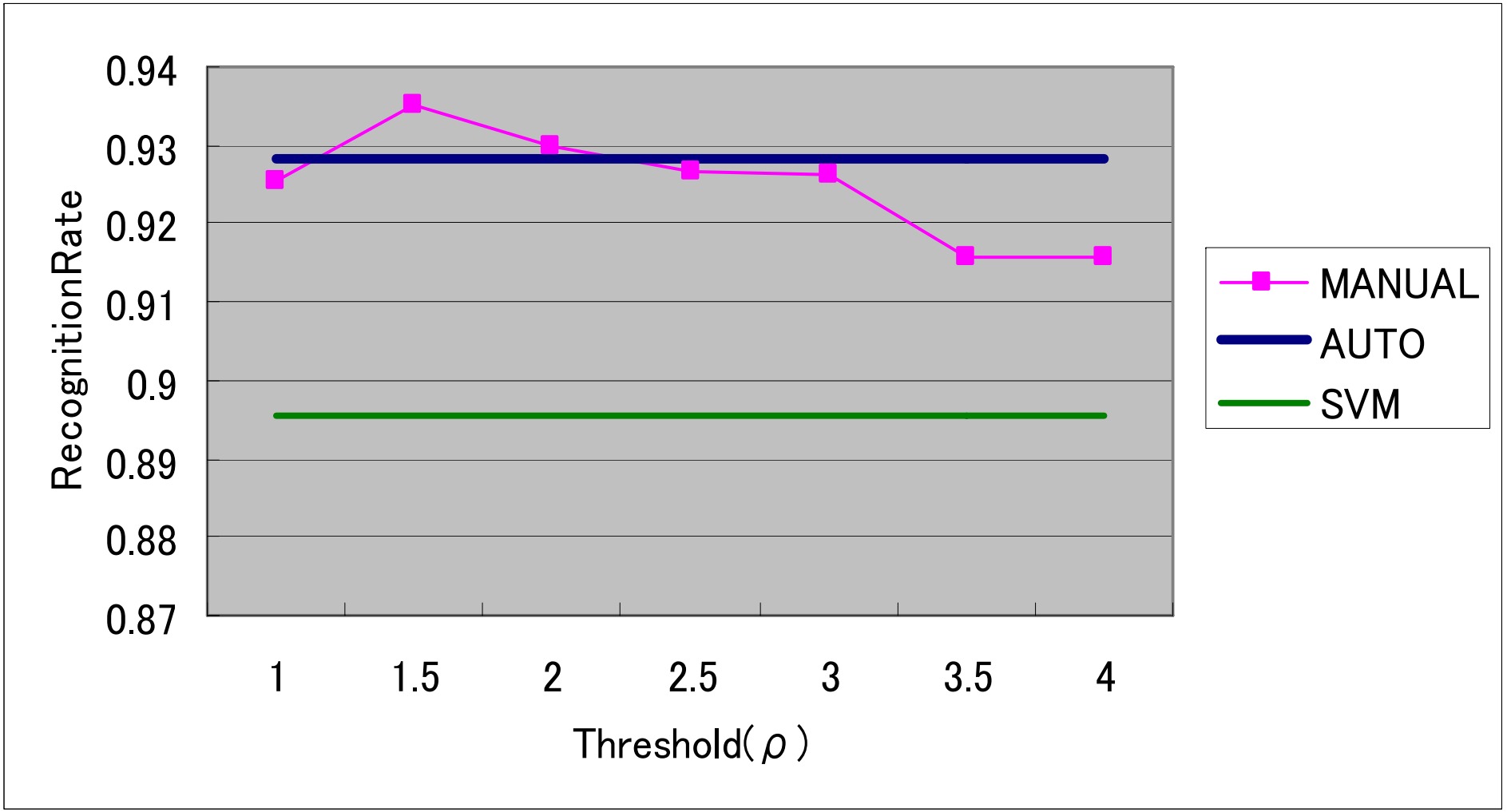
Comparison of recognition rates for the detection time of 60 min



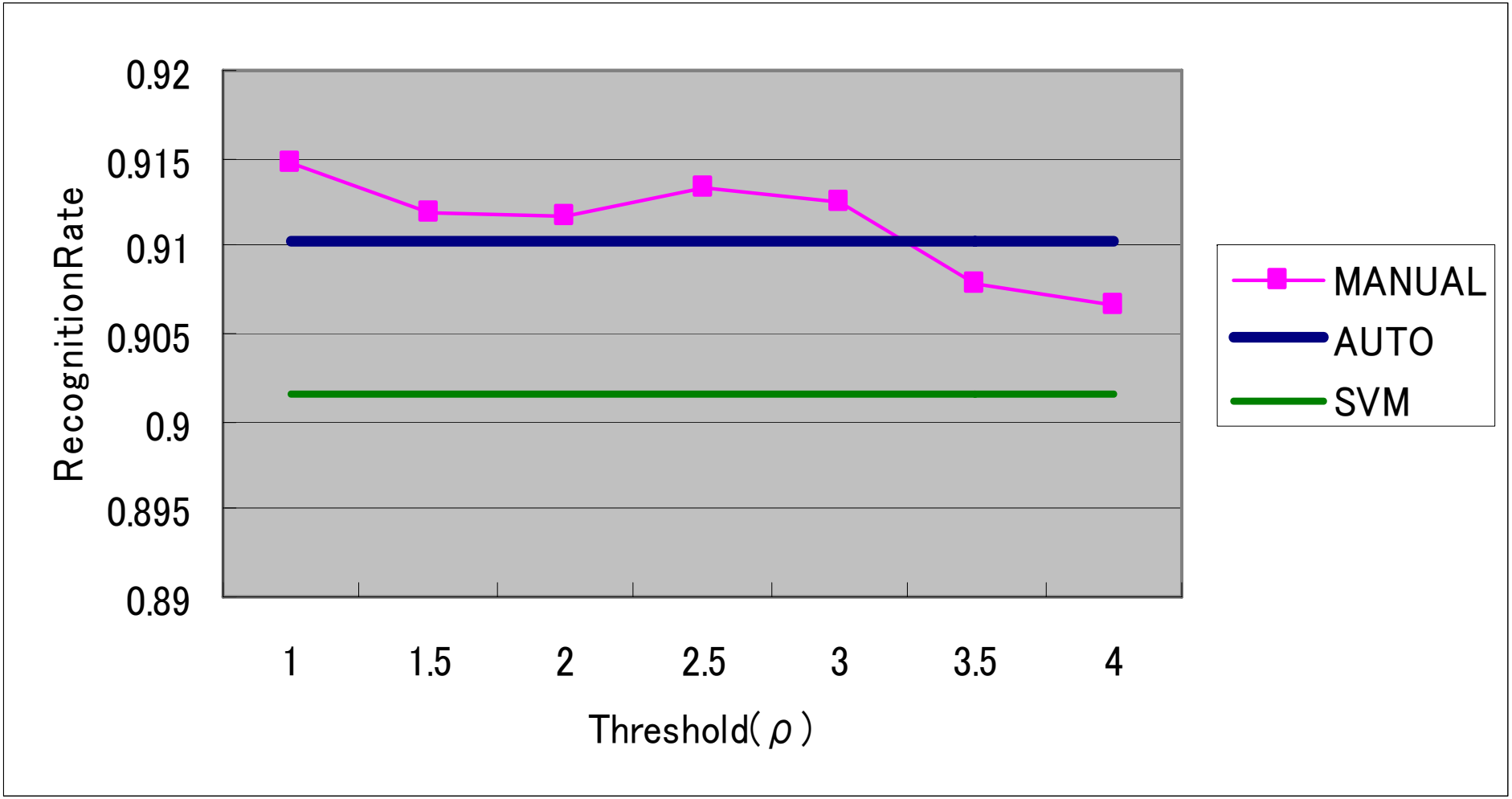
Comparison of recognition rates for the detection time of 45 min



Comparison of recognition rates for the detection time of 30 min



Comparison of recognition rates for the detection time of 15 min



Conclusions and Future Work

- Conclusions
  - BOTs and Humans are classifiable in a short time in high accuracy

- Future Work
  - Labeling Human is costly

Adopt Semi-Supervised Learning using
small # of Humans
medium # of BOTs
very large # of unlabelled characters