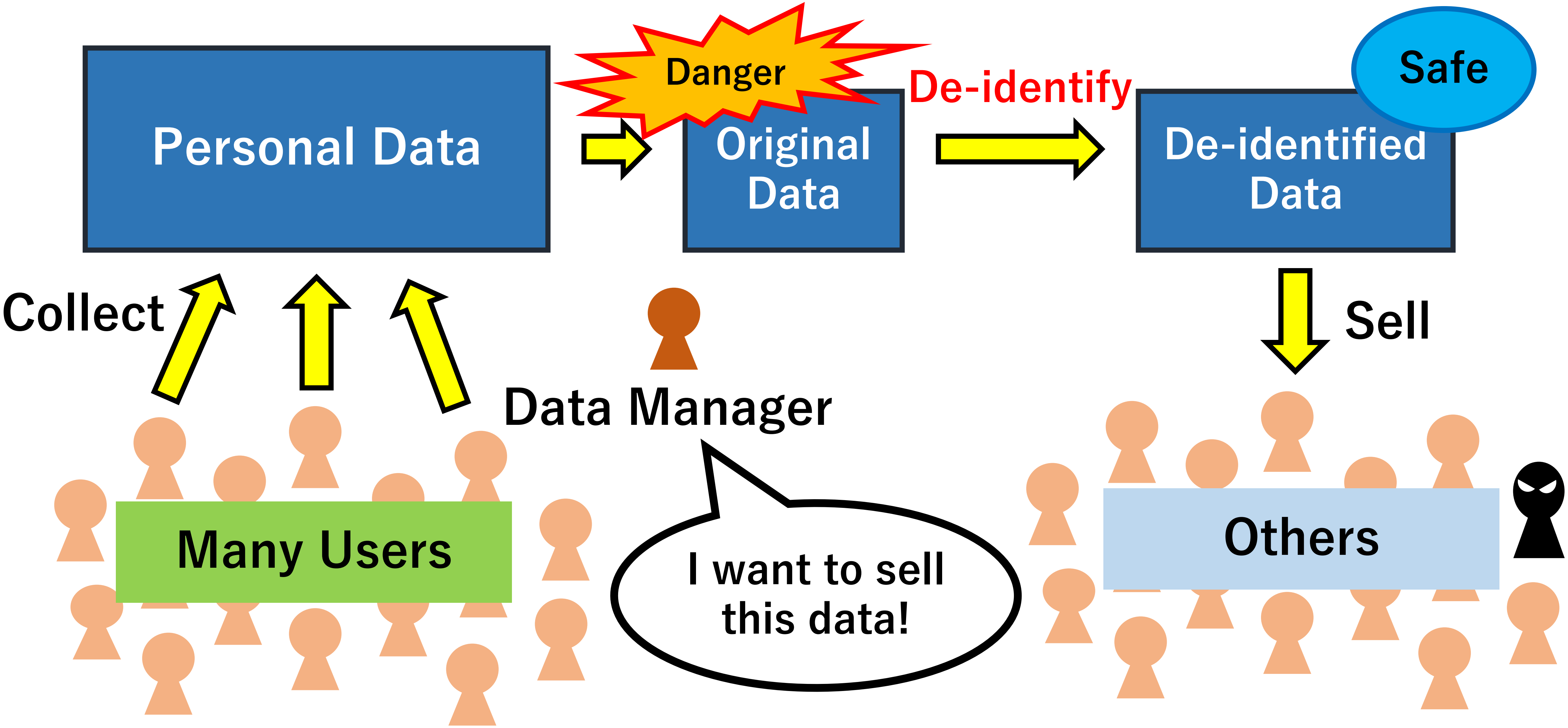
MDAI 2018

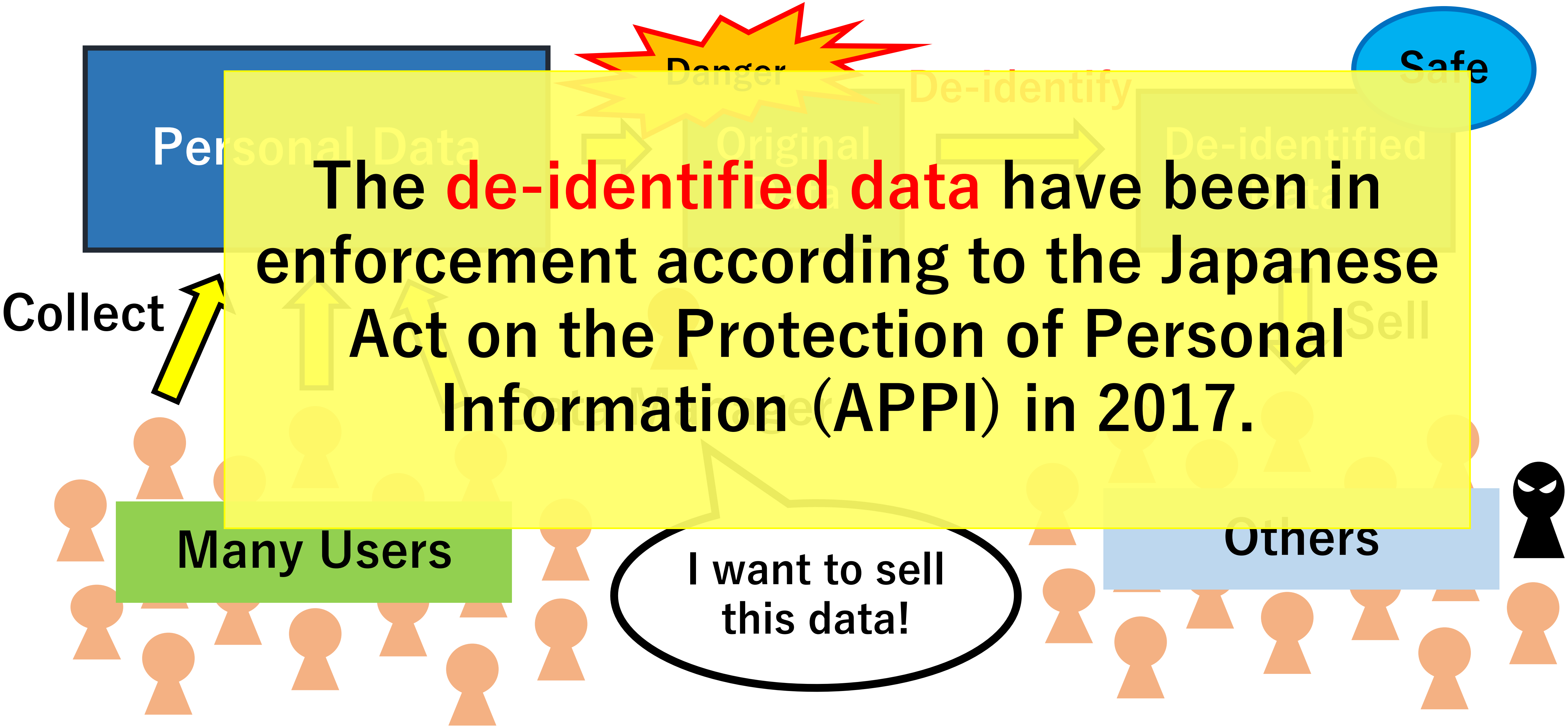
Attacker Models with a Variety of Background Knowledge of Payment History

Satoshi Ito, Hiroaki Kikuchi (Meiji University)
Hiroshi Nakagawa
(RIKEN Center for Advanced Intelligence Project)

What is De-identification?



What is De-identification?

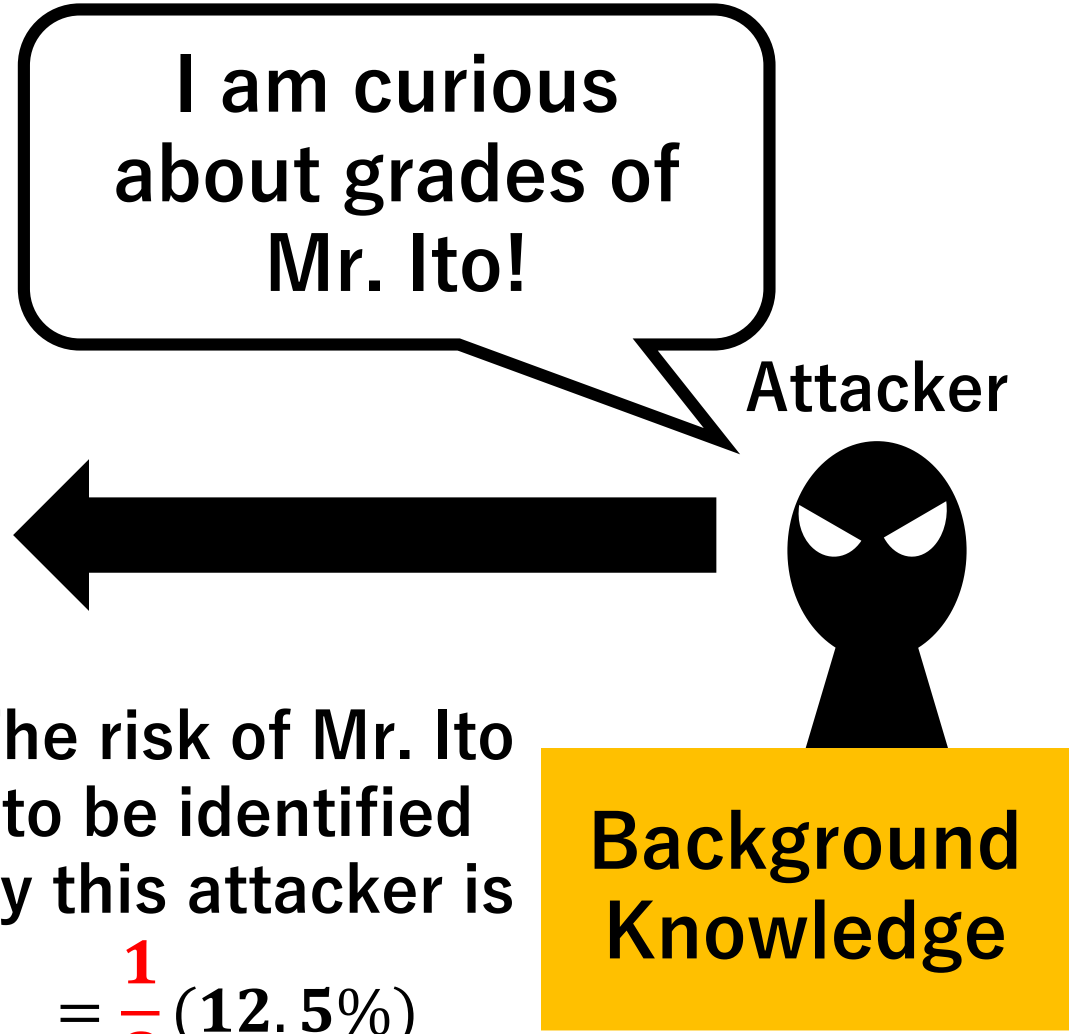


The de-identified data have been in enforcement according to the Japanese Act on the Protection of Personal Information (APPI) in 2017.

Attacker and Background Knowledge

Exam Results

ID	Math	English	Physics
A	90	50	70
B	90	50	60
C	90	70	70
D	50	70	60
E	50	50	80
F	50	50	10
G	30	70	80
H	30	70	10



The risk of Mr. Ito to be identified by this attacker is

$$= \frac{1}{8}(12.5\%)$$

Attacker and Background Knowledge

Exam Results

ID	Math	English	Physics
A	90	50	70
B	90	50	60
C	90	70	70
D	50	70	60
E	50	50	80
F	50	50	10
G	30	70	80
H	30	70	10

Attacker α

Mr. Ito’s English grade must be 50

The risk of Mr. Ito by attacker X

$$= \frac{1}{4}(25\%)$$

Attacker β

Mr. Ito’s Physics grade must be 10

The risk of Mr. Ito by attacker Y

$$= \frac{1}{2}(50\%)$$

Attacker and Background Knowledge

Exam Results

ID	Math	English	Physics
A	90	50	70
B	90	50	60
C	90	70	70
D	50	70	60
E	50	50	80
F	50	50	10
G	30	70	80
H	30	70	10

Attacker α

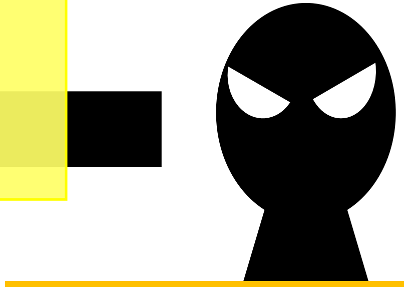
Mr. Ito's English grade must be 50

The risk of Mr. Ito by attacker X

$= \frac{1}{4} (25\%)$

The risk of data depends on the Attacker's background knowledge

Attacker β



Mr. Ito's Physics grade must be 10

The risk of Mr. Ito by attacker Y

$= \frac{1}{2} (50\%)$

Research Question

- **What kind of background knowledge is risky?**
- **Which attribute is the riskiest in data?**

Solution

- **We propose a theoretical risk model which allows to quantify risk without developing re-identification programs.**

Sample Data

Transaction sample data of 4 users in 3 days

ID	User ID	Receipt ID	Date	Time	Goods	Price	Number
1	A	1	2010/12/1	8:45	Apple	1	10
2	C	2	2010/12/1	10:20	Cup	1	30
3	D	3	2010/12/1	16:40	Book	10	5
4	B	4	2010/12/2	9:00	Apple	2	50
5	C	4	2010/12/2	10:00	Book	100	2
6	D	4	2010/12/2	20:00	Cup	20	5
7	A	5	2010/12/3	6:10	Apple	1	10
8	B	6	2010/12/3	10:00	Book	5	5
9	D	7	2010/12/3	12:20	Cup	50	1

9
records

Sample Data

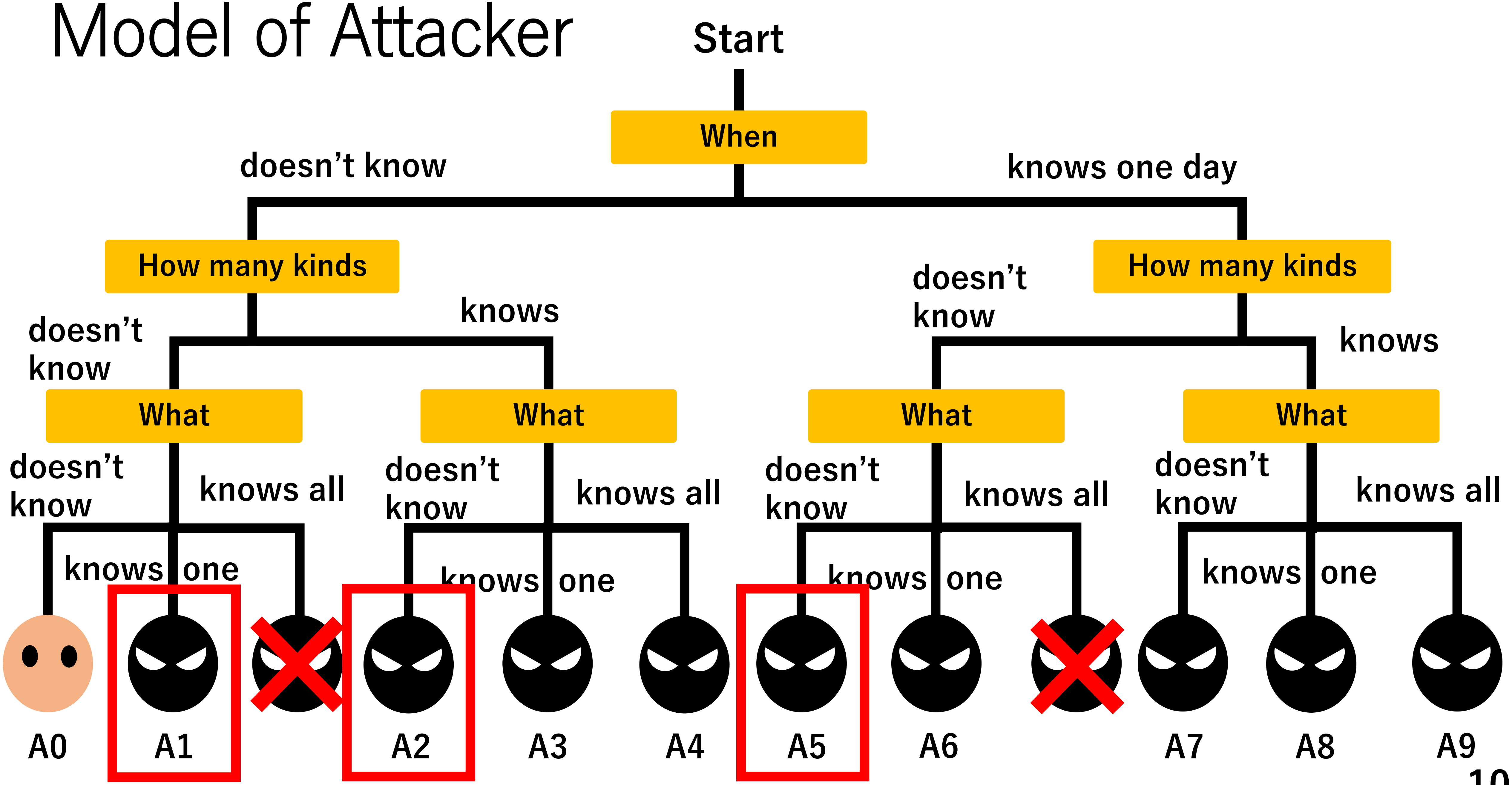
Transaction sample data of 4 users in 3 days

ID	User ID	Receipt ID	Date	Time	Goods	Price	Number
1	A	1	2010/12/1	8:45	Apple	1	10
2	C	2	2010/12/1	10:20	Cup	1	30
3	D	3	2010/12/1	16:40	Book	10	5
4	B	4	2010/12/2	9:00	Apple	2	50
5	C	[illegible]	[illegible]	[illegible]	Book	[illegible]	2
6	D	4	2010/12/2	20:00	Cup	20	5
7	A	5	2010/12/3	6:10	Apple	1	10
8	B	6	2010/12/3	10:00	Book	5	5
9	D	7	2010/12/3	12:20	Cup	50	1

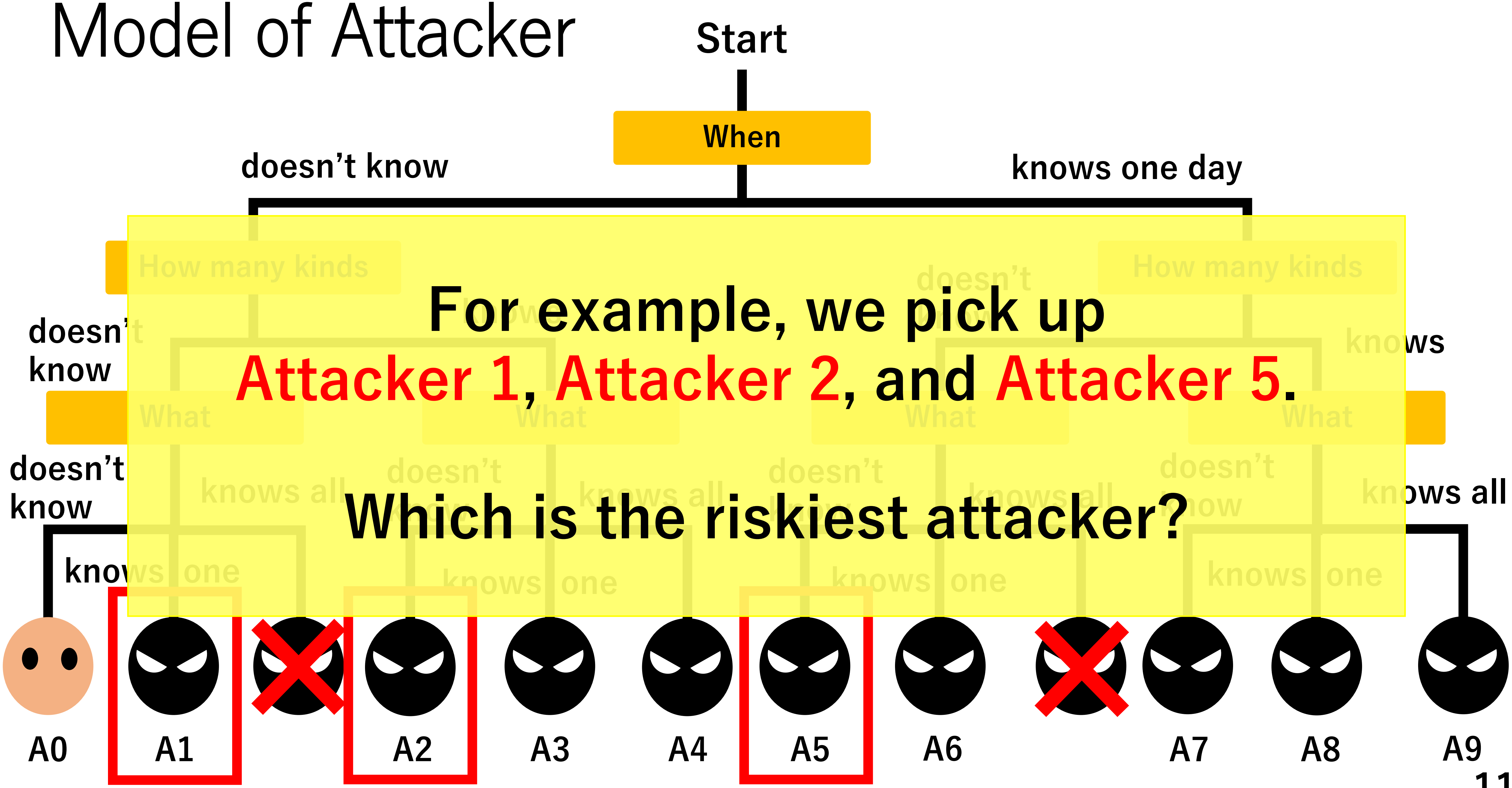
1. **When did he/she buy?**
2. **How many kinds did he/she buy?**
3. **What did he/she buy?**

9 records

Model of Attacker



Model of Attacker



For example, we pick up **Attacker 1**, **Attacker 2**, and **Attacker 5**.

Which is the riskiest attacker?

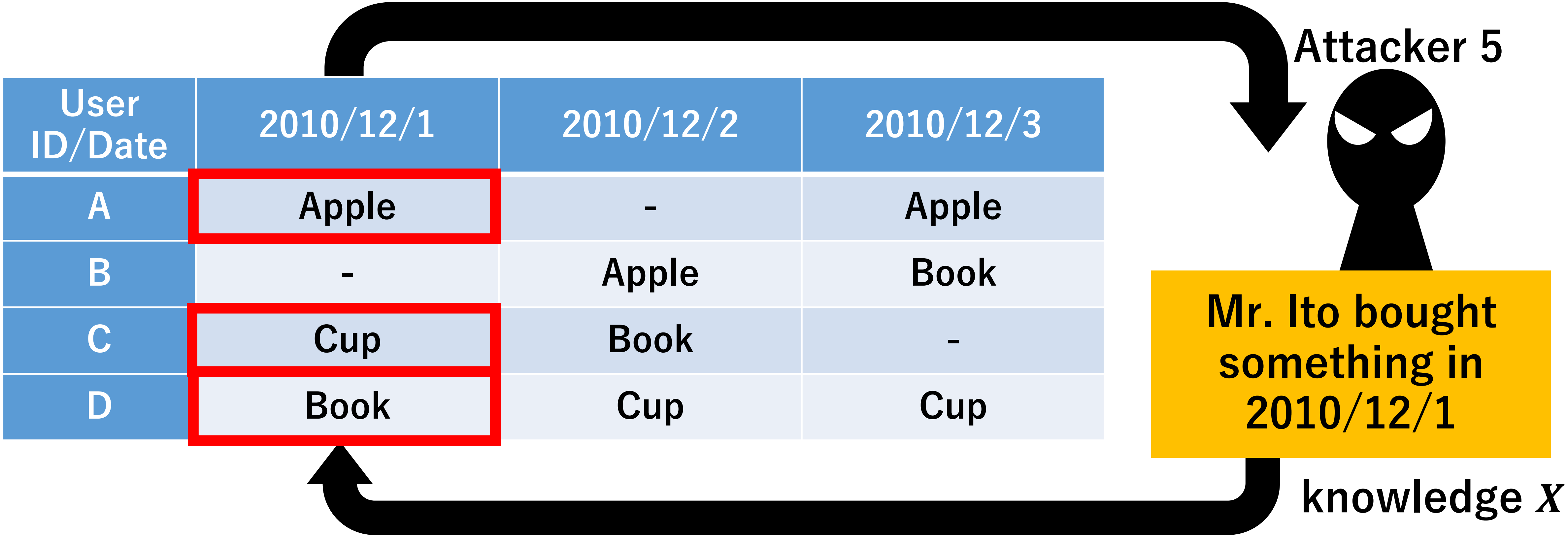
Transformed Sample Data

ID	User ID	Receipt ID	Date	Time	Goods	Price	Number
1	A	1	2010/12/1	8:45	Apple	1	10
2	C	2	2010/12/1	10:20	Cup	1	30
…	…	…	…	…	…	…	…

User ID/Date	2010/12/1	2010/12/2	2010/12/3
A	Apple	-	Apple
B	-	Apple	Book
C	Cup	Book	-
D	Book	Cup	Cup

Risk of Attacker 5

Attacker 5 obtains X with in probability of $\frac{3}{9}$



User ID/Date	2010/12/1	2010/12/2	2010/12/3
A	Apple	-	Apple
B	-	Apple	Book
C	Cup	Book	-
D	Book	Cup	Cup

Mr. Ito bought something in 2010/12/1

knowledge X

Attacker 5 identifies Mr. Ito in probability of $\frac{1}{3}$

Risk of Attacker 5 in this case

$$= \frac{3}{9} \cdot \frac{1}{3} = \frac{1}{9}$$

Mean Identification Probability $\boldsymbol{Pr}(\mathbf{identify}, \boldsymbol{X})$

Attacker 5

Ito bought something in 2010/12/1

knowledge X_1

Attacker 5

Ito bought something in 2010/12/2

knowledge X_2

Attacker 5

Ito bought something in 2010/12/3

knowledge X_3

Mean id.prob,
Risk of Attacker 5
$Pr(\mathrm{identify}, X)$

$$\frac{\mathbf{1}}{\mathbf{9}} + \frac{\mathbf{1}}{\mathbf{9}} + \frac{\mathbf{1}}{\mathbf{9}} = \frac{\mathbf{1}}{\mathbf{3}}$$

Mean Identification Probability $\boldsymbol{Pr}(\mathbf{identify}, \boldsymbol{X})$

Attacker 5

Attacker 5

Attacker 5

Attacker 5 identifies individual with mean probability of $\frac{1}{3}$ when he obtains background knowledge.

Ito bought something in 2010/12/1

Ito bought something in 2010/12/2

Ito bought something in 2010/12/3

Mean id.prob, Risk of Attacker 5

knowledge X_1

knowledge X_2

knowledge X_3

$Pr(\text{identify}, X)$

$$\frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{1}{3}$$

Assumption 1 for Modeling

X: an element of the set of background knowledge.

R_X: set of records that satisfy X

U_X : set of users that satisfy X

Assumption 1: $|R_X| = |U_X|$

$$R_X = \{1, 2, 3\}$$
$$U_X = \{A, C, D\}$$
$$|R_X| = |U_X| = 3$$

Transaction sample data of 4 users in 3 days

ID	User ID	Receipt ID	Date	Time	Goods	Price	Number
1	A	1	2010/12/1	8:45	Apple	1	10
2	C	2	2010/12/1	10:20	Cup	1	30
3	D	3	2010/12/1	16:40	Book	10	5
4	B	4	2010/12/2	9:00	Apple	2	50
…	…	…	…	…	…	…	…

Modeling of Risk of Attackers

$\boldsymbol{m}$: number of records
$\boldsymbol{X}$: an element of the set of background knowledge $\boldsymbol{D(X)}$.
$\boldsymbol{\omega_X = |D(X)|}$

Theorem 4.1

When $|U_X| = |R_X|$, the mean identification probability is

$$Pr(\text{attacked with } X) = \sum_{X \in D(X)} \frac{1}{|U_X|} \frac{|R_X|}{m} = \frac{\omega_X}{m}$$

Assumption 2 for Modeling

$p(X)$: the probability of gaining background knowledge X
$p(Y)$: the probability of gaining background knowledge Y

Assumption 2: $p(X, Y) = p(X)p(Y)$
(X and Y are independent)

Example: X = "2010/12/1“, Y = "$Apple$"

Goods ID /Date	2010/12/1	2010/12/2	2010/12/3
Apple	1	1	1
Book	1	1	1
Cup	1	1	1

Assumption 2 for Modeling

$p(X)$: the probability of gaining background knowledge X

$p(Y)$: the probability of gaining background knowledge Y

Assumption 2: $p(X, Y) = p(X)p(Y)$

(X and Y are independent)

$$\boldsymbol{p(X = \text{"2010/12/1"}) = \frac{1}{3},\quad p(Y = \text{"100"}) = \frac{1}{3}}$$

$$\boldsymbol{p(X = \text{"2010/12/1"})p(Y = \text{"100"}) = \frac{1}{9}}$$

$$\boldsymbol{= p(X = \text{"2010/12/1"}, Y = \text{"100"})}$$

Example: $\boldsymbol{X = \text{"2010/12/1“}, Y = \text{"}Apple\text{"}}$

Goods ID /Date	2010/12/1	2010/12/2	2010/12/3
Apple	1	1	1
Book	1	1	1
Cup	1	1	1

Modeling of Risk of Attackers

$\boldsymbol{m}$: number of records

$\boldsymbol{X}$,$\boldsymbol{Y}$: an element of the set of background knowledge $\boldsymbol{D}(\boldsymbol{X})$, $\boldsymbol{D}(\boldsymbol{Y})$ in table $\boldsymbol{T}$.

$\boldsymbol{\omega_X} = |\boldsymbol{D}(\boldsymbol{X})|$, $\boldsymbol{\omega_Y} = |\boldsymbol{D}(\boldsymbol{Y})|$

Theorem 4.2

When assumption 1, 2 are satisfied, the mean identification probability is

$$\boldsymbol{Pr}(\textbf{attacked with } \boldsymbol{X}, \boldsymbol{Y}) = \frac{\boldsymbol{\omega_X \omega_Y}}{\boldsymbol{m}}$$

Actual value and Accuracy of Our Model

ID/date	2010/12/1	2010/12/2	2010/12/3
A	Apple	-	Apple
B	-	Apple	Book
C	Cup	Book	-
D	Book	Cup	Cup

Attacker 5

$$\mathbf{Pr(attacked\ with\ } \boldsymbol{date}\mathbf{)} = \frac{\boldsymbol{\omega_{date}}}{\boldsymbol{m}} = \frac{\mathbf{3}}{\mathbf{9}} = \color{red}{\frac{\mathbf{1}}{\mathbf{3}}}$$

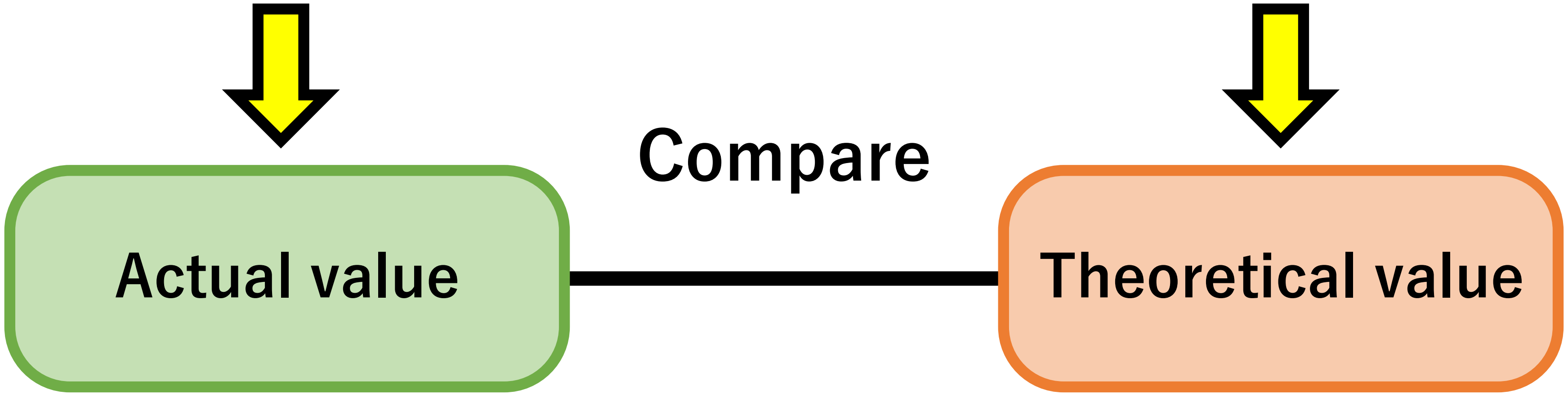
$$\text{Acutual value} = \frac{\mathbf{1}}{\mathbf{9}} + \frac{\mathbf{1}}{\mathbf{9}} + \frac{\mathbf{1}}{\mathbf{9}} = \color{red}{\frac{\mathbf{1}}{\mathbf{3}}}$$

Evaluation of Our Model

Transaction data of 400 users in 1 year

User ID	Receipt ID	Date	Time	Goods ID	Price（＄）	Num
12583	536370	2010/12/1	8:45	22728	3.75	24
12583	536370	2010/12/1	8:45	22727	3.75	24
12583	536370	2010/12/1	8:45	22726	3.75	12
12583	536370	2010/12/1	8:45	21724	0.85	12
…	…	…	…	…	…	…

38087 records



Compare

Actual value

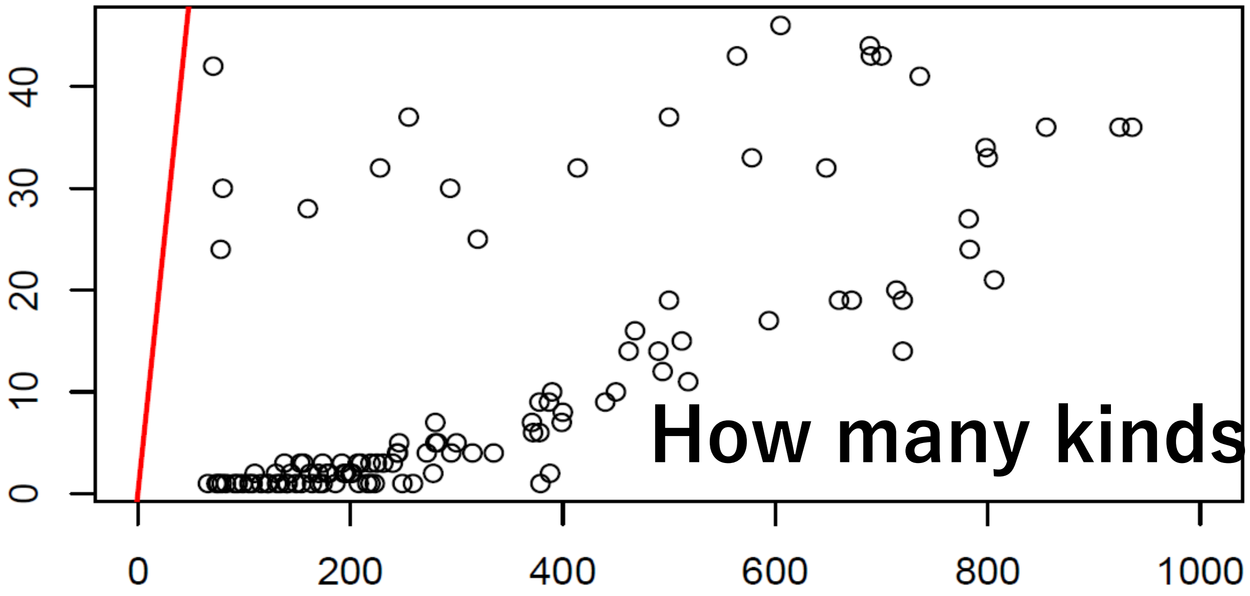
Theoretical value

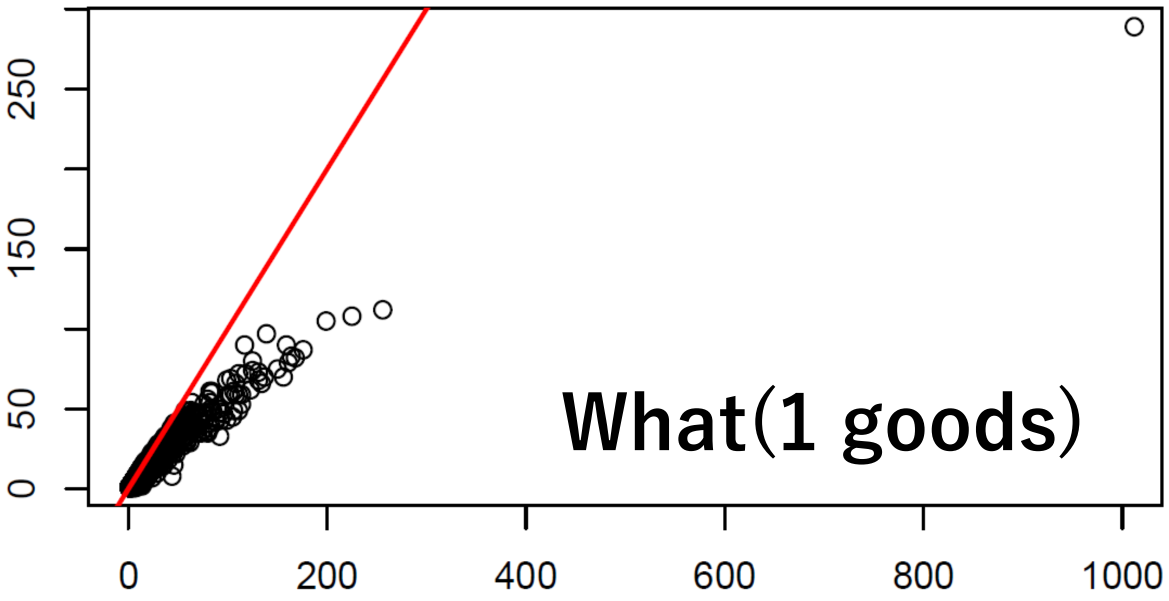
Experimental Results

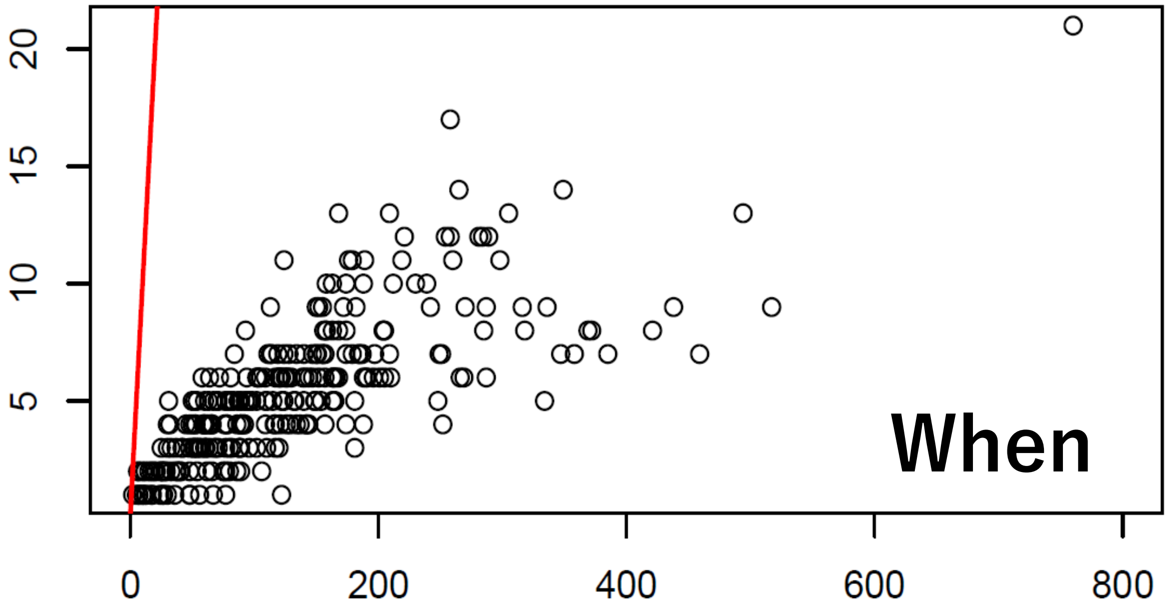
Attacker	Actual value	Theoretical value	When	How many kinds	What
0	0.0025	0.0025	-	-	-
1	0.0965	0.0730	-	-	One
2	0.0807	0.0030	-	✓	-
3	0.7974	8.3240	-	✓	One
4	0.9788	4.5440	-	✓	All
5	0.1851	0.0076	✓	-	-
6	0.8945	21.1700	✓	-	One
7	0.9400	0.8680	✓	✓	-
8	0.9750	2415.0000	✓	✓	One
9	0.9994	1319.0000	✓	✓	All

Discussion

Scatter plot of $|R_X|$ and $|U_X|$
x-axis: $|R_X|$, y-axis: $|U_X|$
Red Line: $|R_X| = |U_X|$





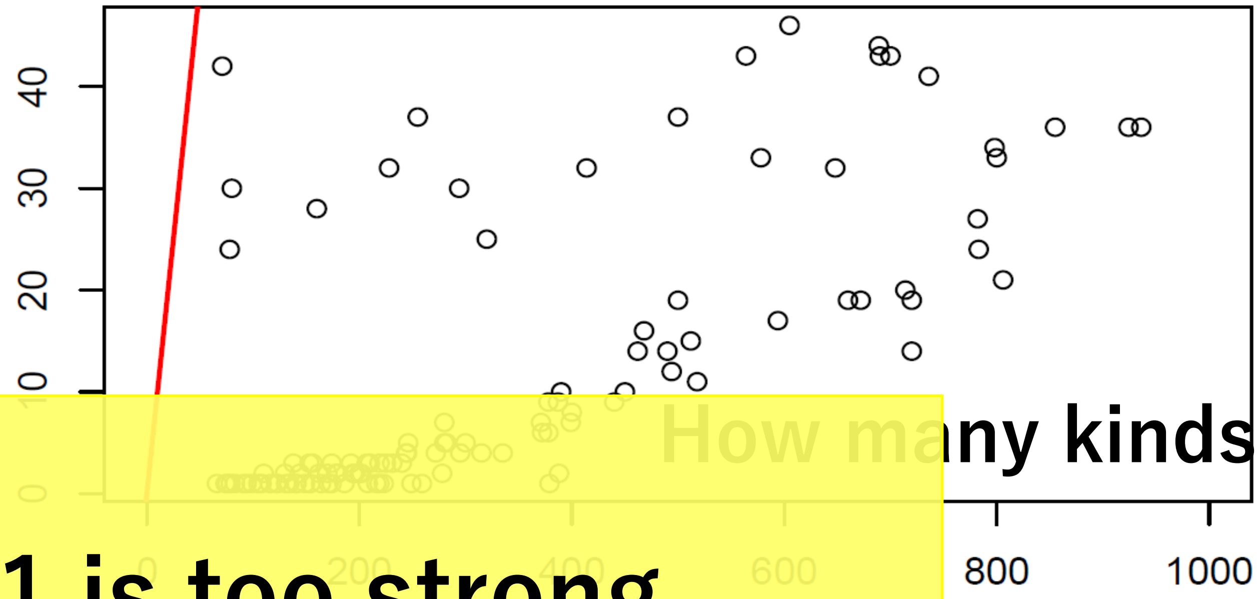


Discussion

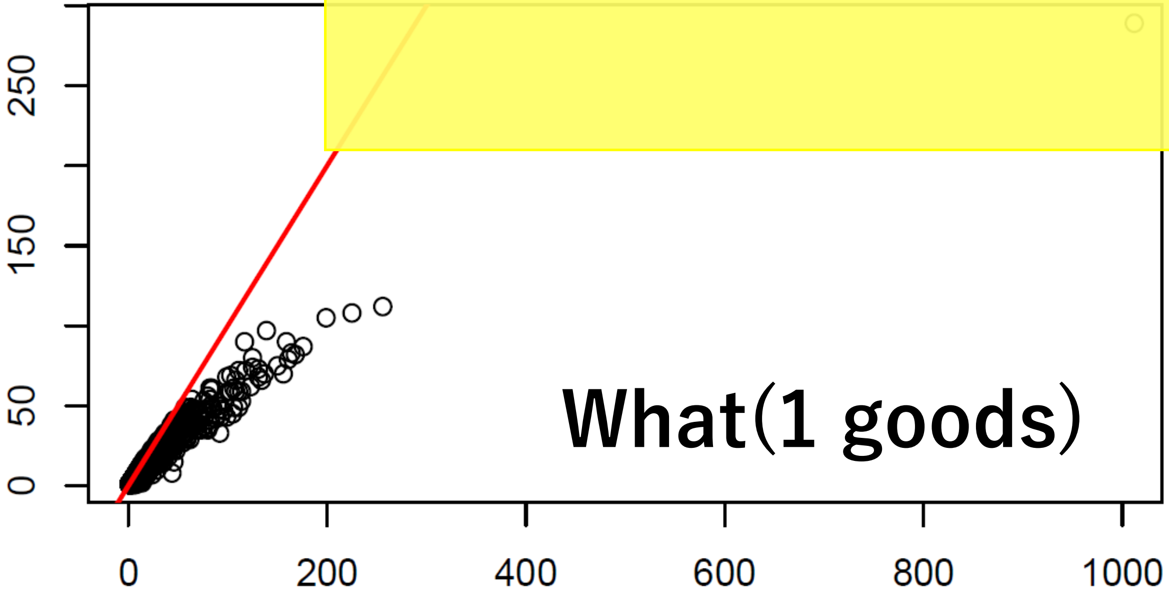
Scatter plot of $|R_X|$ and $|U_X|$

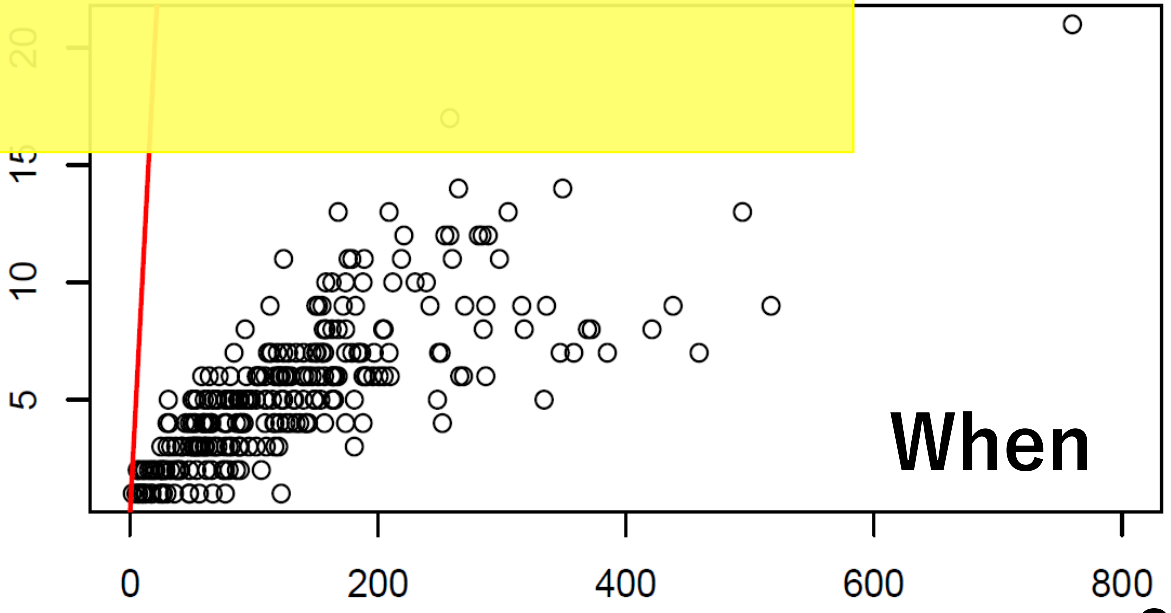
x-axis: $|R_X|$, y-axis: $|U_X|$

Red Line: $|R_X| = |U_X|$



Assumption 1 is too strong.





Conclusions

- **We proposed 10 types of attackers with background knowledge about 400 and evaluated the risk (mean identification probability) associated with these attackers.**

- **We found that date is the most useful for attackers among three kinds of background knowledge: purchase date, number of kinds, and knowledge of one good purchased.**

- **We demonstrated that the risk can be theoretically estimated without computing it exactly under two assumptions.**