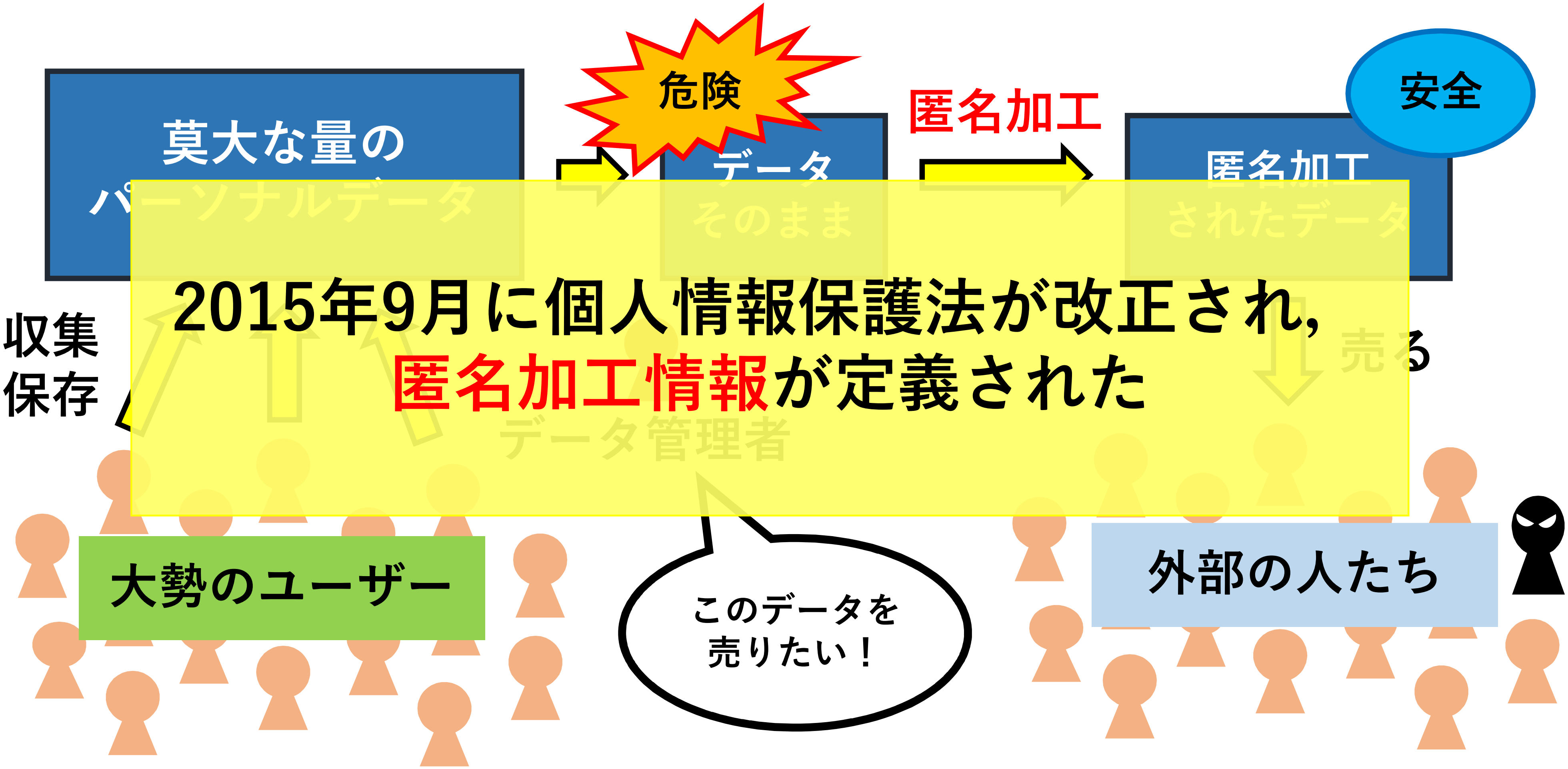
PWS 2017

背景知識の違いによる匿名加工データの攻撃者モデルの分類と評価

伊藤聡志，菊池浩明（明治大学）

中川裕志（東京大学）

匿名加工とは？



2015年9月に個人情報保護法が改正され,
匿名加工情報が定義された

攻撃者と背景知識

菊池研学生の試験結果

ID	数学	英語	物理
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

どれかが伊藤

このデータから
伊藤の試験結果を
知りたい！

攻撃者

攻撃者が
このデータから
伊藤を識別
できる確率

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

背景知識

攻撃者と背景知識

菊池研学生の試験結果

ID	数学	英語	物理
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

攻撃者の持つ背景知識によってデータの危険度は変わる

攻撃者X

伊藤の英語の点数は50点である

伊藤を識別できる確率 $= \frac{1}{4}(25\%)$

攻撃者Y

伊藤の物理の点数は10点である

伊藤を識別できる確率 $= \frac{1}{2}(50\%)$

本研究について

研究目的

- どんな背景知識を持つ攻撃者が危険なのか？
- データ中のどの属性が危険であるのか？

困難性

- データの安全性を測るためにはアルゴリズムを設計し，再識別プログラムを開発する必要がある

本研究について

PWSCUP2017で用いられている安全性指標

RE-IDENTIFY	S1－datenum	野島 良 Ryo Nojima	匿名加工前後で，購入日と数量の組み合わせが等しいレコード同士を同じ顧客とみなす. If the 2nd （date） and the 6th （number） attributes are the same between some anonymized record i and some non－anonymized record j， then the algorithm regards i as the anonymized record of j and outputs customer id in j
RE-IDENTIFY	S2－[illegible]	野島 良 [illegible]	匿名加工前後で，商品IDと単価の組み合わせが等しいレコード同士を同じ顧客とみなす. If the 4th （gift） and the 5th （price） attributes are the same between some anonymized [illegible]
RE-IDENTIFY	S3－[illegible]	野島 良 [illegible]	匿名加工前後で，商品IDと数量の組み合わせが等しいレコード同士を同じ顧客とみなす. [illegible] and the 6th （number） [illegible] between some [illegible] the algorithm regards i as [illegible]
RE-IDENTIFY	S4－itemdate	[illegible]	匿名加工前後で，[illegible]レコード同士を同じ顧客とみなす. [illegible] are the same between some [illegible] record j， then the algorithm regards i as the anonymized record of j and outputs customer id in j
RE-IDENTIFY	S5－ item2pricenum	濱田 浩気 Koki Hamada	匿名加工前後で，商品ID2桁，価格と個数の組み合わせが等しいレコード同士を同じ顧客とみなす. If the 4th （gift id）， the 5th （price） and the 6th （number） attributes are the same between some anonymized record i and some non－anonymized record j， then the algorithm regards i as the anonymized record of j and outputs customer id in j
RE-IDENTIFY	S6－ item2datenum	濱田 浩気 Koki Hamada	匿名加工前後で，商品ID2桁，購入日と個数の組み合わせが等しいレコード同士を同じ顧客とみなす. If the 4th （gift id）， the 2nd （date） and the 6th （number） attributes are the same between some anonymized record i and some non－anonymized record j， then the algorithm regards i as the anonymized record of j and outputs customer id in j

プログラムの開発や実行を必要としないで
データの再識別リスクを評価する
数理モデルを提案する

トイデータ

4人分の3日間の購買履歴データ(例)

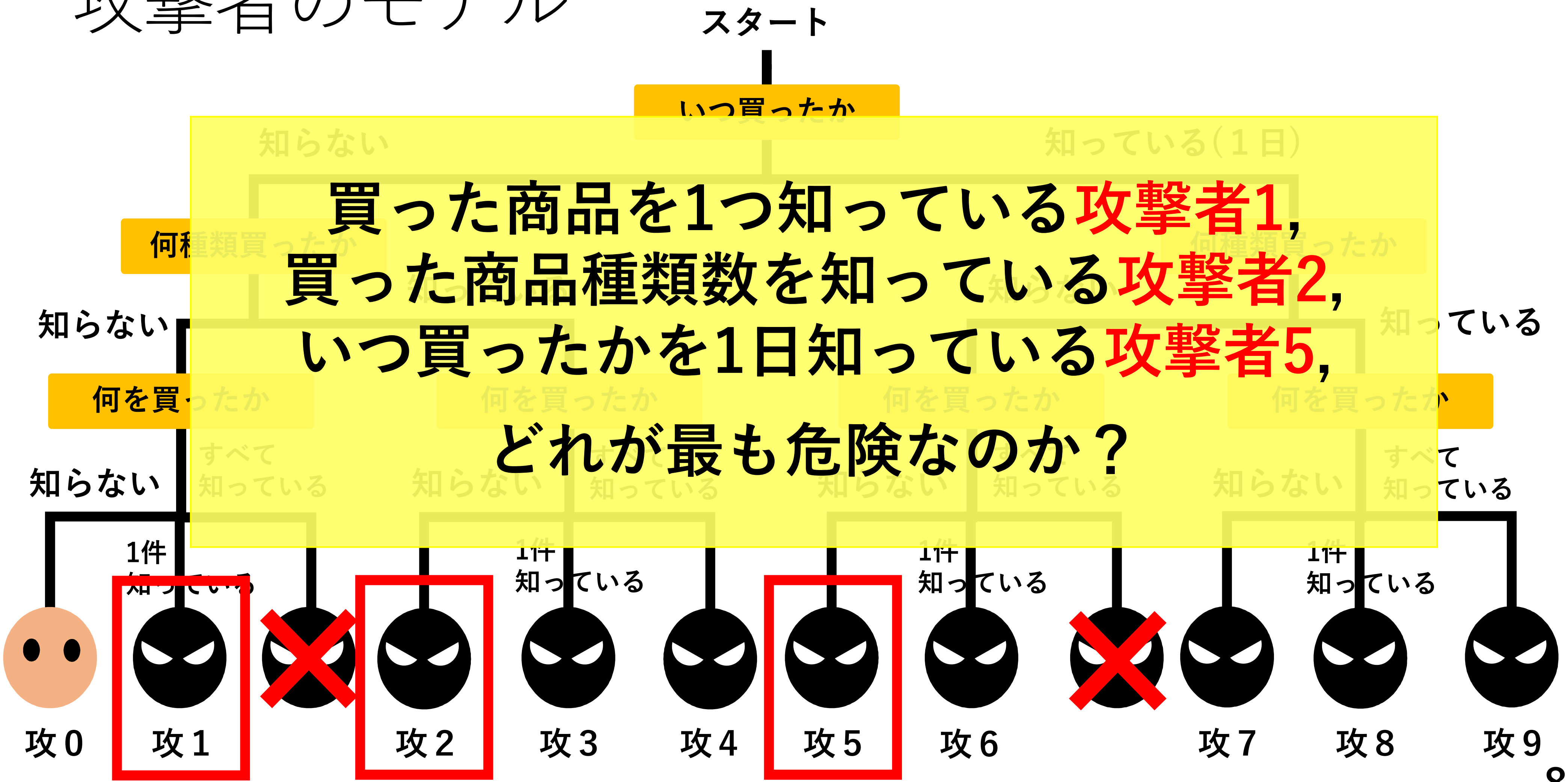
レコード	ID	購買ID	年月日	時	商品ID	単価($)	個数
1	A	1	2010/12/1	8:45	100	1	10
2	C	2	2010/12/1	10:20	100	1	30
3	D	[illegible]	2010/12/1	16:40	200	10	5
4	B	4	2010/12/2	9:00	100	2	50
5	B	[illegible]	2010/12/2	9:00	200	100	2
6	[illegible]	[illegible]	2010/12/2	9:00	300	20	5
7	A	5	2010/12/3	6:10	100	1	10
8	B	6	2010/12/3	10:00	200	5	5
9	D	7	2010/12/3	12:20	100	50	1
10	D	8	2010/12/3	20:00	300	1	100

10行

1. いつ買ったか
2. 何種類買ったか
3. 何を買ったか

に注目する

攻撃者のモデル



買った商品を1つ知っている**攻撃者1**,
買った商品種類数を知っている**攻撃者2**,
いつ買ったかを1日知っている**攻撃者5**,

どれが最も危険なのか？

トイデータを変換した表

レコード	ID	購買ID	年月日	時	商品ID	単価($)	個数
1	A	1	2010/12/1	8:45	100	1	10
2	C	2	2010/12/1	10:20	100	1	30
…	…	…	…	…	…	…	…

⇩

ID/購入日	2010/12/1	2010/12/2	2010/12/3
A	100		100
B		100, 200, 300	200
C	100		
D	200		100, 300

攻撃者 5 の危険度

攻撃者5がこの背景知識を得る確率

$$Pr(X = "2010/12/1") = \frac{3}{10}$$

購買履歴データを変換したもの

ID/購入日	2010/12/1	2010/12/2	2010/12/3
A	100		100
B		100, 200, 300	200
C	100		
D	200		100, 300

攻撃者5

伊藤は
2010/12/1に
買い物をした

背景知識 X

攻撃者5が伊藤を識別できる確率

$$Pr(id|X = "2010/12/1") = \frac{1}{3}$$

この場合の攻撃者の危険度

$$= \frac{3}{10} * \frac{1}{3} = \frac{1}{10}$$

平均識別確率 $Pr(\text{identify}|X)$

攻撃者5

攻撃者5

攻撃者5

攻撃者5は背景知識を手に入れたとき

平均 $\frac{8}{15}$ の確率で個人を識別できる

ある顧客は
2010/12/1に
買い物をした

背景知識 x_1

ある顧客は
2010/12/2に
買い物をした

背景知識 x_2

ある顧客は
2010/12/3に
買い物をした

背景知識 x_3

このデータに対する
攻撃者5の危険度
$Pr(\text{identify}|X)$

$$\frac{3}{10} * \frac{1}{3} + \frac{3}{10} * \frac{1}{1} + \frac{4}{10} * \frac{1}{3} = \frac{8}{15}$$

数理モデルのための仮定１

$\boldsymbol{R_x}$: 背景知識 x に該当するレコード行の集合

$\boldsymbol{U_x}$: 背景知識 x に該当する顧客の集合

仮定1： $|R_x| = |U_x|$

例：背景知識 x = 「2010/12/1に買い物をした」

ID/購入日	2010/12/1	2010/12/2	2010/12/3
A	100		100
B		100, 200, 300	200
C	100		
D	200		100, 300

$$R_x = \{1, 2, 3\}$$
$$U_x = \{A, C, D\}$$
$$|R_x| = |U_x| = 3$$

危険度の数理モデル化

$\boldsymbol{m}$：レコード数　$\boldsymbol{X}$：背景知識のカテゴリ

$\boldsymbol{\omega_X}$：背景知識の種類数

例）$\boldsymbol{\omega}_{\text{購買日}} = \mathbf{3}\ (2010/12/1, 2010/12/2, 2010/12/3)$

定理4.1

仮定1のもと，単一の背景知識 $\boldsymbol{X}$ を持つ攻撃者の平均識別確率は,

$$\boldsymbol{Pr}(\mathbf{identify}|\boldsymbol{X}) = \sum_{x \in X} \frac{\mathbf{1}}{|\boldsymbol{U_x}|} \frac{|\boldsymbol{R_x}|}{\boldsymbol{m}} = \frac{\boldsymbol{\omega_X}}{\boldsymbol{m}}$$

である.

数理モデルのための仮定 2

$Pr(X)$: 背景知識 X が起きる確率
$Pr(Y)$: 背景知識 Y が起きる確率

仮定2： $Pr(X,Y) = Pr(X)Pr(Y)$

例：背景知識 x = 「2010/12/1に買い物をした」
背景知識 y = 「商品100を買った」

商品ID /購入日	2010/12/1	2010/12/2	2010/12/3
100	2	1	2
200	0	1	1
300	1	1	1

$$Pr(X=x) = \frac{3}{10}, Pr(Y=y) = \frac{5}{10}$$

$$Pr(X=x)Pr(Y=y) = \frac{15}{100} \approx \frac{2}{10}$$

危険度の数理モデル化

$\boldsymbol{m}$：レコード数　　$\boldsymbol{X, Y}$：背景知識のカテゴリ

$\boldsymbol{\omega_X, \omega_Y}$：背景知識の種類数

定理4.2

仮定1のもと，単一の背景知識 $\boldsymbol{X}$ と独立な $\boldsymbol{Y}$ を同時に持つ攻撃者の平均識別確率は,

$$\boldsymbol{Pr}(\mathbf{identify}|\boldsymbol{X, Y}) = \frac{\boldsymbol{\omega_X \omega_Y}}{\boldsymbol{m}}$$

である.

危険度の理論値と実測値

ID/購入日	2010/12/1	2010/12/2	2010/12/3
A	100		100
B		100, 200, 300	200
C	100		
D	200		100, 300

攻撃者5

$$\mathbf{Pr(identify|購買日)} = \frac{\omega_{購買日}}{m} = \frac{\mathbf{3}}{\mathbf{10}} = \mathbf{0.3}$$

$$\textbf{実際に計算した危険度} = \frac{\mathbf{8}}{\mathbf{15}} = \mathbf{0.53}$$

実験：数理モデル評価

400人分の1年間の購買履歴データ

ID	購買ID	年月日	時	商品ID	単価($)	個数
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行

攻撃者0～9の危険度の実測値

危険度の大小関係を比較する

攻撃者0～9の危険度の理論値

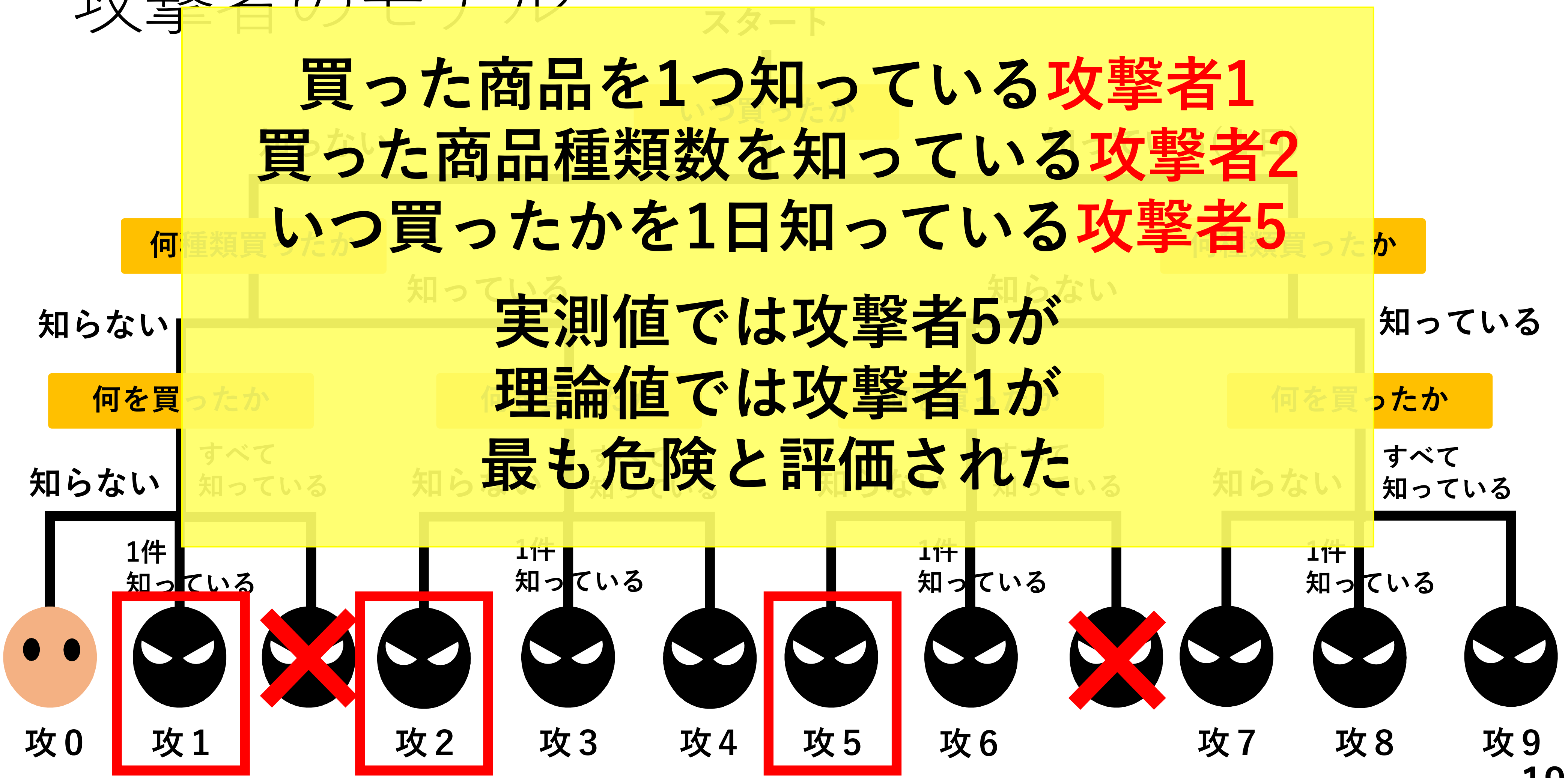
実験結果

攻撃者	危険度 （実測値）	危険度 （理論値）	いつ	何種類	何を
0	0.0025	0.0025	×	×	×
1	0.0965	0.0730	×	×	1商品
2	0.0807	0.0030	×	○	×
3	0.7974	8.3240	×	○	1商品
4	0.9788	4.5440	×	○	全商品
5	0.1851	0.0076	○	×	×
6	0.8945	21.1700	○	×	1商品
7	0.9400	0.8680	○	○	×
8	0.9750	2415.0000	○	○	1商品
9	0.9994	1319.0000	○	○	全商品

攻撃者のモデル

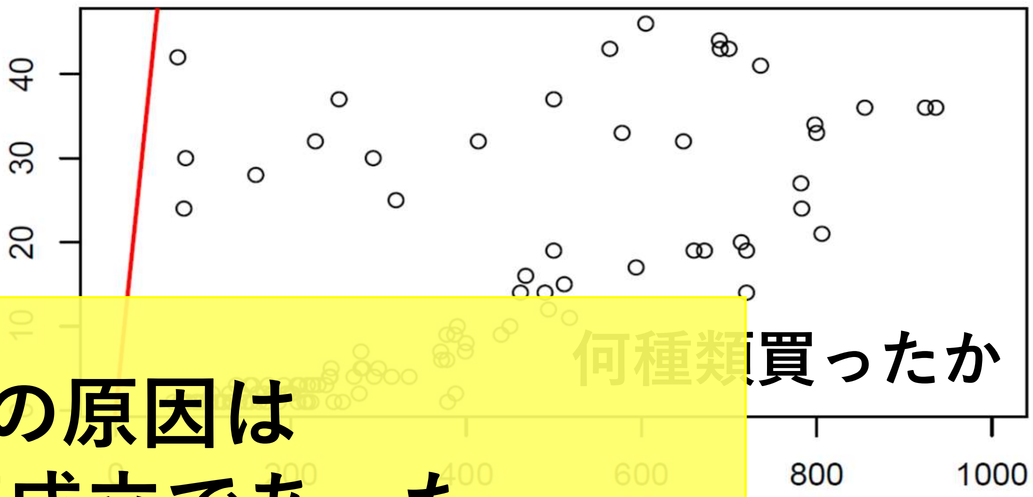
買った商品を1つ知っている**攻撃者1**
買った商品種類数を知っている**攻撃者2**
いつ買ったかを1日知っている**攻撃者5**

実測値では攻撃者5が
理論値では攻撃者1が
最も危険と評価された

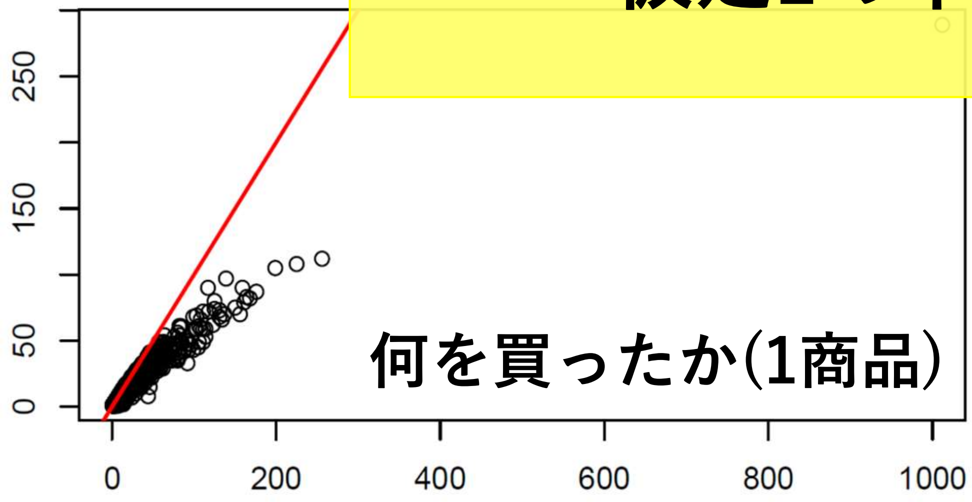


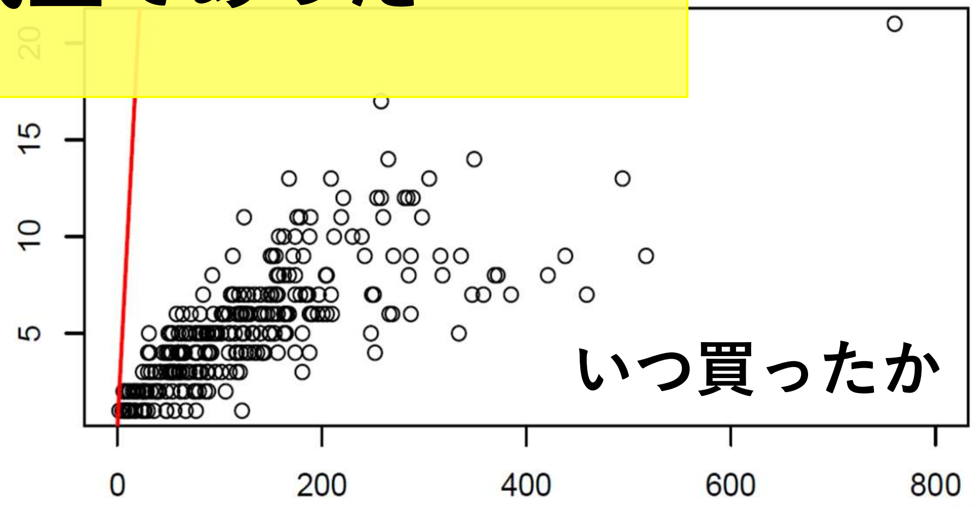
考察

$|R_x|$と$|U_x|$の散布図
横軸：$|R_x|$，縦軸：$|U_x|$
赤い直線：$|R_x| = |U_x|$



誤差の原因は
仮定1の不成立であった





まとめ

- 「いつ買ったか？」「何を買ったか？」「何種類買ったか？」に注目し，背景知識の異なる10タイプの攻撃者を想定した.
- 攻撃者の危険度を仮定を置いて数理モデル化し，400人分の1年間の購買履歴データを用いて危険度の理論値と実測値を比較した.
- その結果，理論値では「何を買ったか？」，実測値では「いつ買ったか？」を知る攻撃者が危険であると評価された.