

PWSCUP-2019

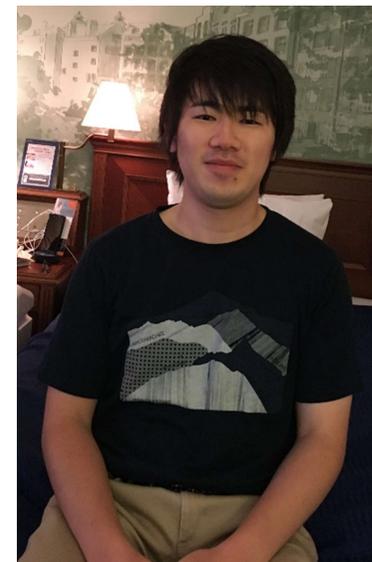


# チーム & 戦略紹介

代表者 & 発表者  
伊藤聡志(明治大学)

# チーム紹介

明治大学（菊池研）と静岡大学（西垣研・大木研）  
による，2大学3研究室の合同チーム



これまでの成績

明治大学

静岡大学

2015

ステテコ伊藤  
ビリから2番目くらい

nifigaki  
総合2位

2016

ステテコ伊藤2  
総合8位・再識別賞

nifigaki  
総合7位

2017

合同チーム：ステテコ西垣  
総合5位

2018

合同チーム：ステテコ西垣2  
総合3位

2019

合同チーム：ステテコ西垣3  
?????

# 加工フェイズ戦略（明大担当）

顧客ごとの滞在地域のばらつきを標準偏差によって評価し、危険な時間帯の地域IDをランダムに変える！

	8:00	8:30	9:00	9:30	10:00	10:30	11:00	11:30	12:00	12:30	13:00	13:30	14:00	14:30	15:00	15:30	16:00	16:30	17:00	17:30
1	92	109	107	108	107	113	115	123	136	128	118	125	129	126	113	126	113	104	101	123
2	64	102	96	111	114	118	127	125	137	130	130	131	132	155	142	162	167	171	167	165
3	8	28	64	57	91	78	101	118	119	117	137	107	118	110	147	148	145	137	141	153
4	25	42	51	58	78	76	87	71	80	93	102	101	99	110	75	111	95	80	90	67
5	19	38	55	73	75	97	93	104	103	101	113	100	106	111	104	115	115	122	118	113
6	27	86	95	117	124	143	137	124	135	161	154	166	169	174	151	148	202	194	193	201
7	60	137	142	140	141	161	161	140	136	98	120	137	122	137	137	134	130	156	167	152
8	55	66	87	101	99	122	122	104	107	113	129	136	138	128	110	93	118	123	107	117
9	46	80	123	133	153	170	170	173	160	164	177	161	151	161	164	152	138	157	157	176

# 識別フェイズ戦略（静大担当）

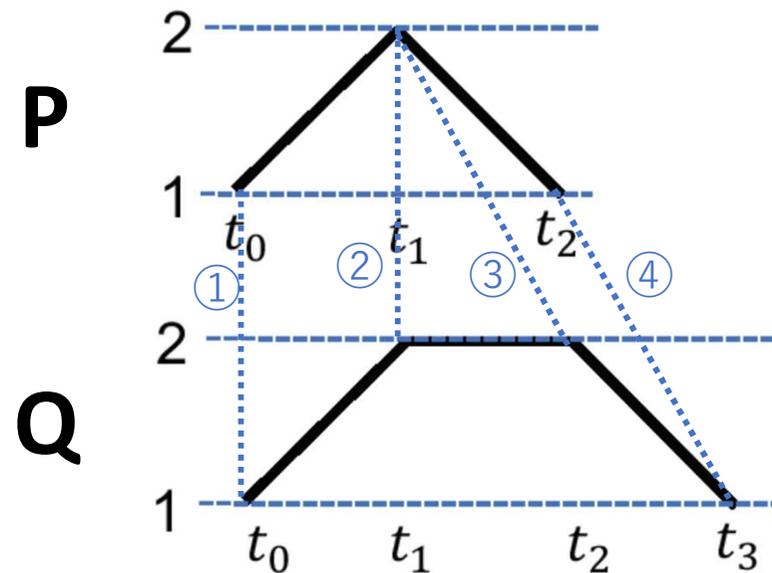
顧客の軌跡間の距離をDTW距離を用いて測定し、個人を識別する！

## DTW距離

2つの時系列データ間の距離を測る指標

トレース推定は

DTW再識別・Jaccard再識別・参照トレース再送攻撃を組み合わせ使っている！

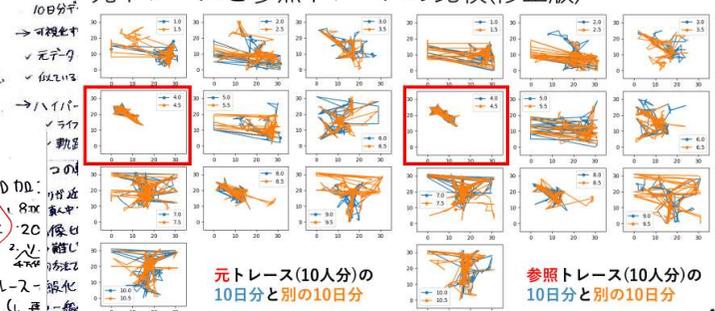


# ポスターの宣伝

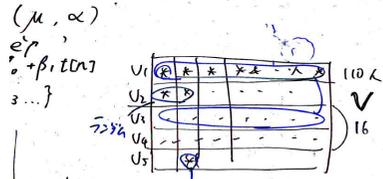
2019  
ステディ会議 10/3

トリス推定  
・再送攻撃以上の物がない  
・ただし攻撃するのはい

## 元トリスと参照トリスの比較(修正版)



4

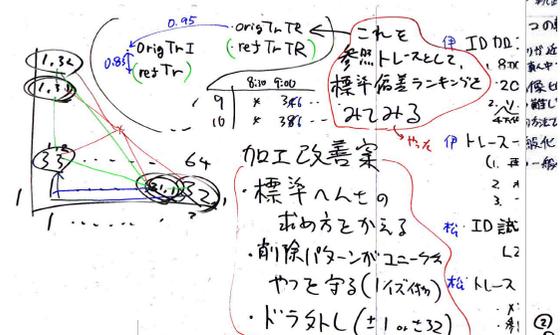


$$\mu = \beta_0$$

$$\tau \sim \text{ga}$$

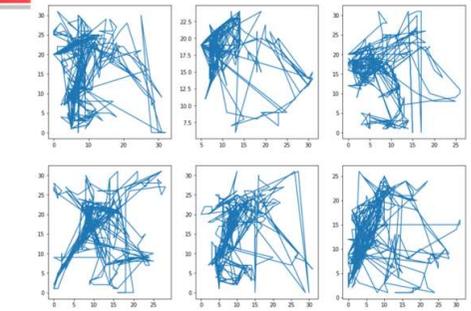
# incident : time CN  
bMS CN  
# pre-incident pre\_time

1. ID 識別  
2. 相違ユーザの追加トリス  
3. ドラ削除  
4. 削除トリスが \* を含む



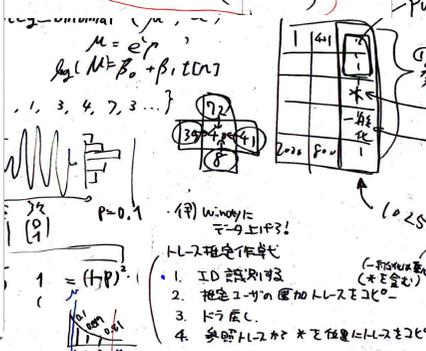
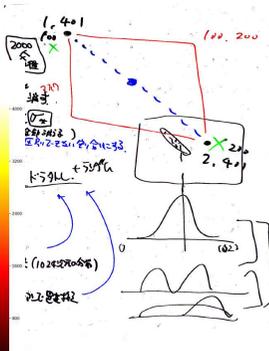
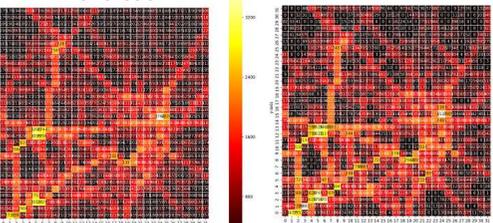
- 加工改善案  
・標準偏差の求め方をかえる  
・前削除パターンがユニークなやつを守る (1/25倍)  
・ドラ外し (±1.032)

## ちょっとやってみた



Nishigaki Laboratory

自チームIDPデータ(20日分) 2000ユーザについて  
加工データ(pubtrace)の所在情報  
公開情報  
自チームIDPデータ(20日分) 2000ユーザについて  
参照データ(reftrace)の所在情報



Q 削除パターン識別は行政に委託  
✓ 削除した1/25倍  
・削除は安全?  
・1/25倍と比較は2/3?  
# 再識別フェーズに  
・ID識別は軽微な問題  
簡単な識別で済む! → 1/25倍  
簡単な再識別の再確認

本戦用データの分析(削除パターン)  
本戦データ(2,000人分)の削除パターンを分析してみた  
削除パターンがユニークな顧客: 397人 識別される可能性アリ

人数が多かった(安全な)削除パターンの例

パターン	8	8.5	9	9.5	10	10.5	11	11.5	12	12.5	13	13.5	14	14.5	15	15.5	16	16.5	17	17.5	人数	
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	230(推)
2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	212
3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	155
4	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	114
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	115
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	106

## ドラの削除(ノイズ付加)

元トリス

user_id	time_id	reg_id
1	401	605
1	402	445
1	403	316
1	404	2
1	405	315
1	406	282
...	...	...
10	800	635

ランダムに ±1か±32する (1マスずらす)

ドラ削除されたトリス

user_id	time_id	reg_id
1	401	605
1	402	445
1	403	316
1	404	34
1	405	315
1	406	282
...	...	...
10	800	635

対トリス推定加工は、ドラ削除+標準偏差低い組み合わせ削除(対ID識別加工)で行おうと考えています!(まだできていない)