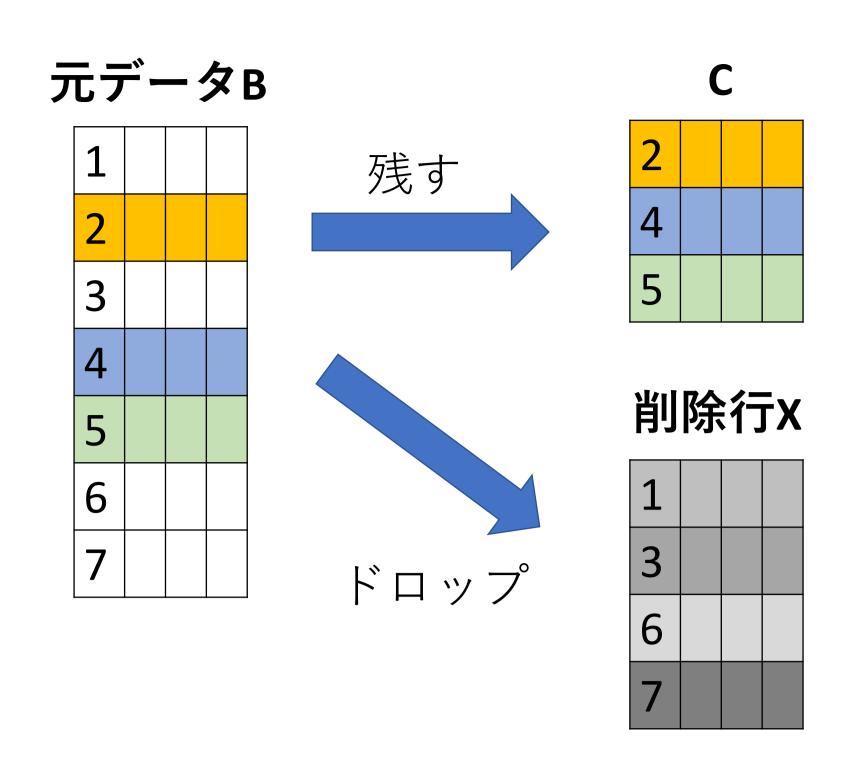
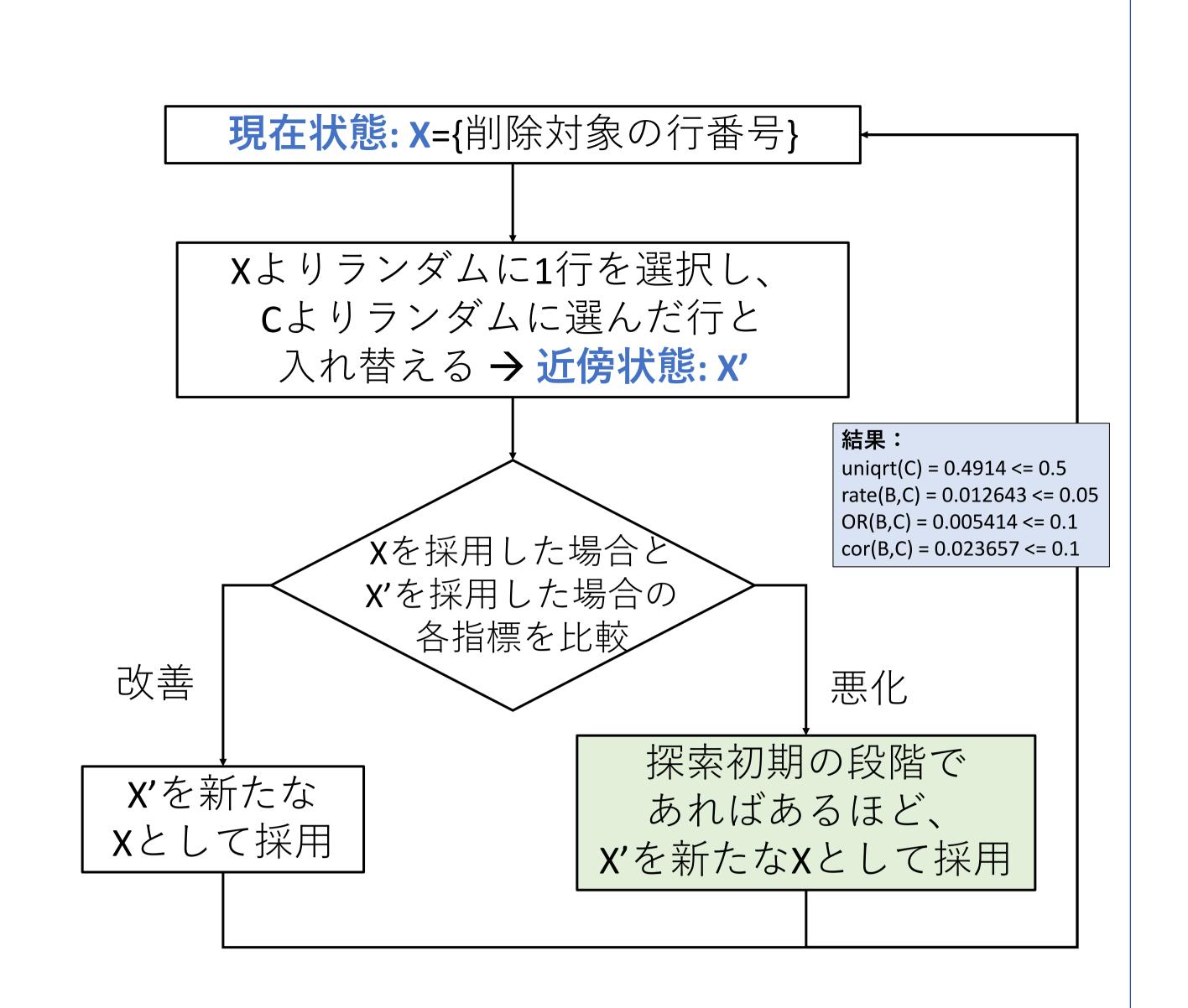
# PWS Cup 2021 チーム「めたぼっち」 (Team #007) 井口誠 (Kii 株式会社)

## 匿名化フェーズ(第1匿名化)

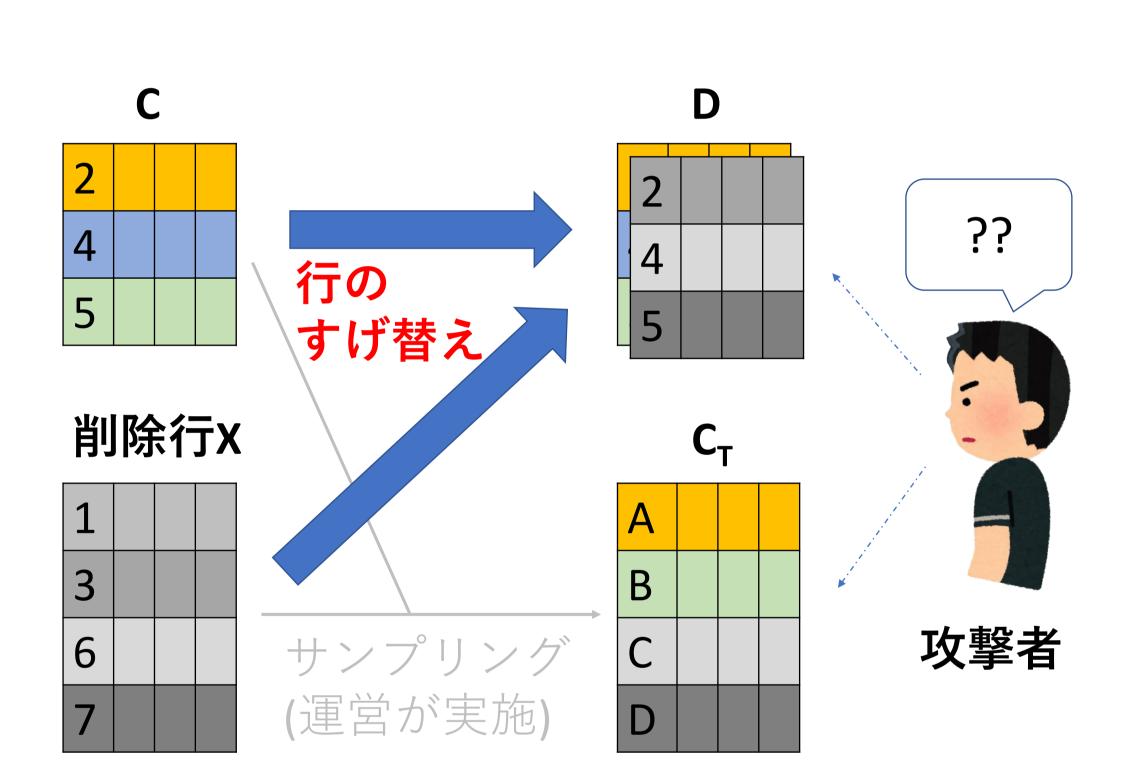


#### 削除行数=1200レコードと固定の上で

安全性指標 uniqrt(C)<= 0.5 を満たしつつ 有用性指標 Umark(B,C)の値ができるだけ 良くなる削除行Xの組み合わせを焼きなまし法 (疑似アニーリング法)で見つける



### 匿名化フェーズ(第2匿名化)



有用性指標 Iloss(C,D) と Umark(B,D) をしきい値内に収めつつ、できるだけ多くのCの行を、削除行Xのいずれかの行とすげ替える組み合わせを焼きなまし法 (疑似アニーリング法)で見つける

#### 正しいしきい値は Iloss (C,D) <=4 現在状態: D Dよりランダムに1行選択し、以下のいずれかを実施 → 近傍状態:D' 高確率で、この行とすげ替え可能な(i..e, lloss <=5 な) 行をXよりランダムに選択し、行のすげ替えを実施 • 低確率で、行を元に戻す(選択行がすげ替え済の場合) 結果: $rate(B,D) = 0.0433 \le 0.05$ $OR(B,D) = 0.0287 \le 0.1$ DとD'の $cor(B,D) = 0.0798 \le 0.1$ 各指標を比較 悪化 改善 探索初期の段階で D'を新たな あればあるほど、 Dとして採用 D'を新たなDとして採用

### 攻撃フェーズ

他のチームが、上記の第1&第2匿名化を使ってデータ加工してきたことを想定して検討を進めた結果...

