



コードクロンの編集過程における 開発者の潜在的な社会構造に関する分析

和歌山大学

©久木田 雄亮, 大平 雅雄

コードクローンとは

- ソースコード中に存在する「同一の」または「類似する」コード片のこと
 - 主に、既存コードを再利用（コピーアンドペースト）することにより作成される

```
public int exeCalcEven(){  
> int counter = 0;  
> for(int i=0; i<10; i++){  
> > if(i%2 == 0){  
> > > counter += i;  
> > > System.out.println("i is " + i);  
> > > System.out.println("counter is " + counter);  
> > }  
> > if(i%2 == 0){  
> > > counter += i;  
> > > System.out.println("i is " + i);  
> > > System.out.println("counter is " + counter);  
> > }  
> }  
> return counter;  
> }
```

コード片A

コード片A'

file i

コードクローン

```
public int exeCalcOdd(){  
> int counter = 0;  
> for(int i=0; i<10; i++){  
> > if(i%2 == 1){  
> > > counter += i;  
> > > System.out.println("i is " + i);  
> > > System.out.println("counter is " + counter);  
> > }  
> }  
> return counter;  
> }
```

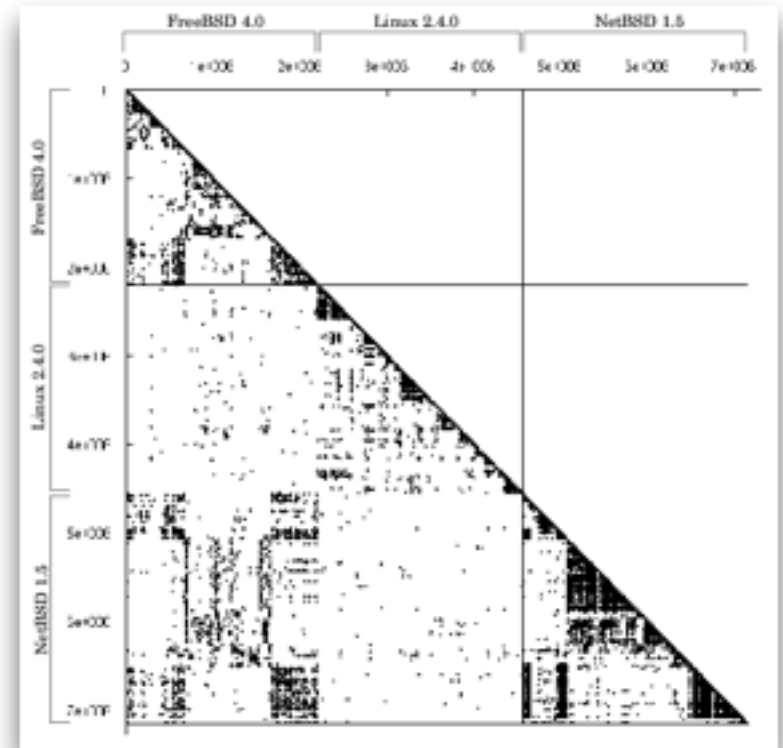
コード片A''

file j

コード片A, A', A''はクローン関係にある

研究の背景 (1/2)

- ソフトウェア開発・保守の過程では、生産性の向上を目的として既存コードが頻繁に再利用されることがある
- コードクローンがソースコード全体に遍在することが少なくない

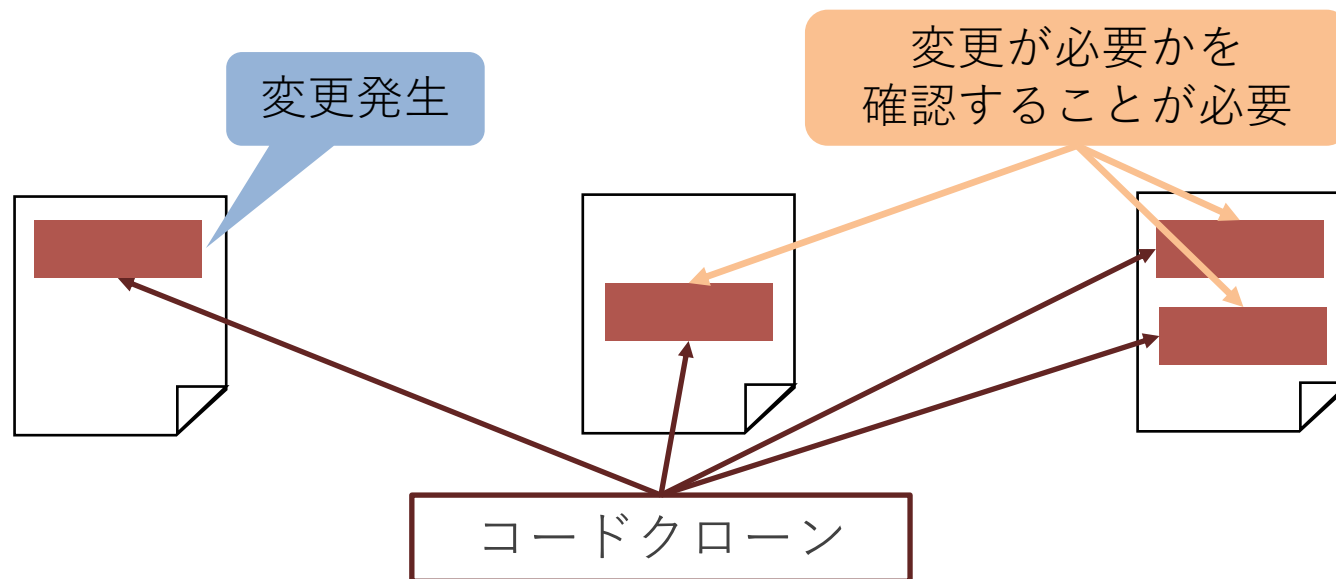


UNIX系OS, Linux間のコードクローン[1]

[1] Kamiya, T., Kusumoto S. and Inoue K.: CCFinder: A multilinguistic token-based code clone detection system for large scale source code, IEEE Transactions on Software Engineering (TSE), Vol.28, No.7, pp.654-670, 2002.

研究の背景 (2/2)

- 欠陥除去などの理由によりコードクローンに変更を加える際のリスク
 - 変更箇所が多岐に渡るため保守コストが増加する
 - 変更漏れにより新たな不具合を混入してしまう



先行研究の問題点

- これまでのコードクローン研究
 - コードクロンの検出手法
 - 主眼：大規模なシステムのソースコードを高速に解析しコードクローンを検出すること
 - コードクロンの分析手法
 - 主眼：コードクローンの特徴と品質（バグ発生確率）との関係を明らかにすること



**コードクロンの作成・利用
過程と品質**の関係に着目した
分析はほとんどおこなわれて
いない

研究の目的

- 目的

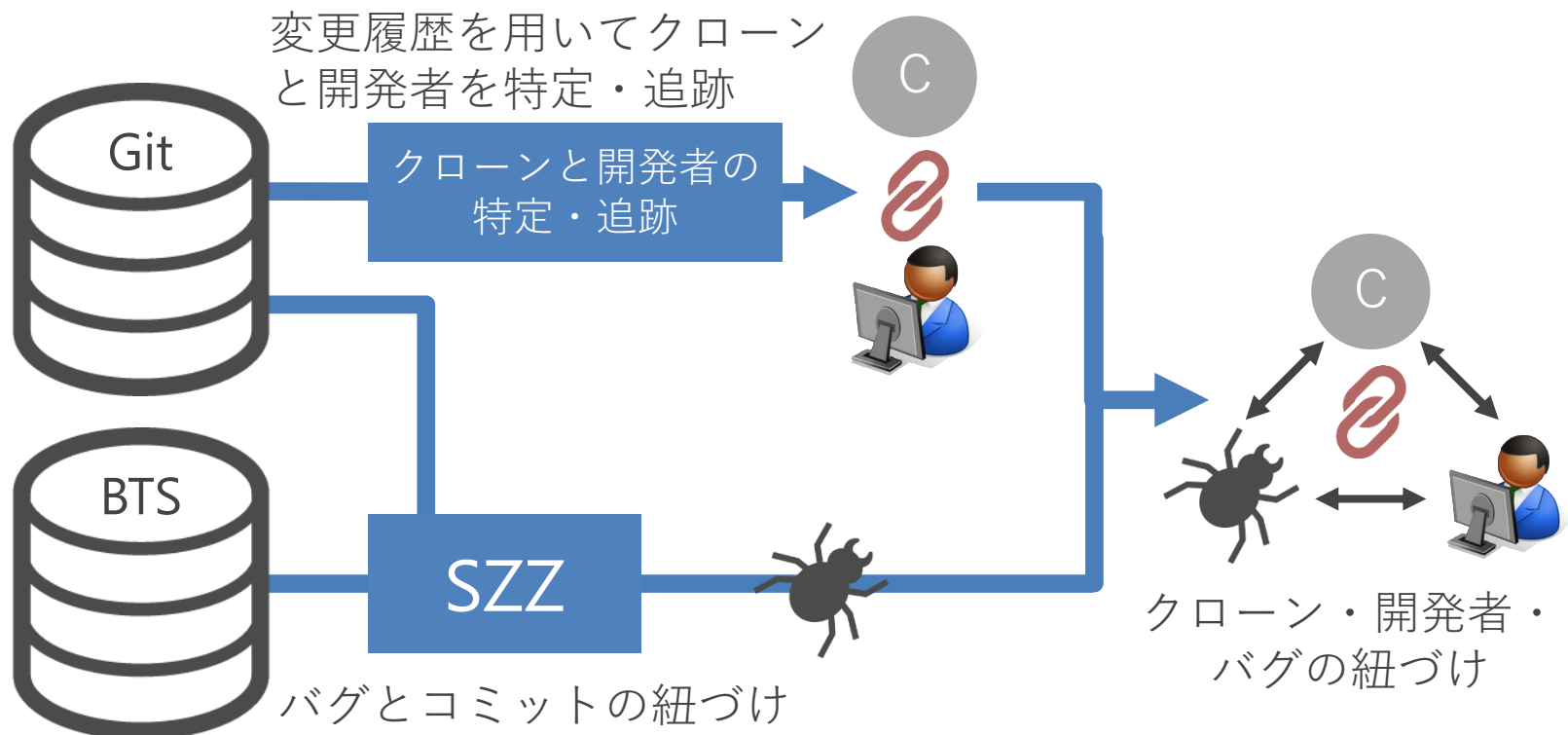
- コードクロンの長期的かつ安全な管理方法について、**コードクロンの作成・利用過程の観点**から有益な知見を提供すること

- アプローチ

- コードクロンの作成・利用過程における人的影響を調査するためのツールを構築する
 - コードクロンの作成・利用過程と開発者の追跡
 - SZZアルゴリズムによるバグとクロンの紐付け
 - コードクロン作成・利用過程と品質の関係の分析

CCT: Code Clone Tracer

- コードクローンの作成・利用過程における人的影響の調査を支援するためツール



Research Questions

- 共同開発を行っている開発者同士とそうでない開発者が作成・変更に関わっているクローンでは品質に違いがあるか
- クローンの変更回数，間隔の大きさがクローンの品質に与える影響に違いがあるか
- クローンを再利用する際，自身が作成したクローンと他者が作成したクローンでクローンの品質に与える影響に違いがあるか