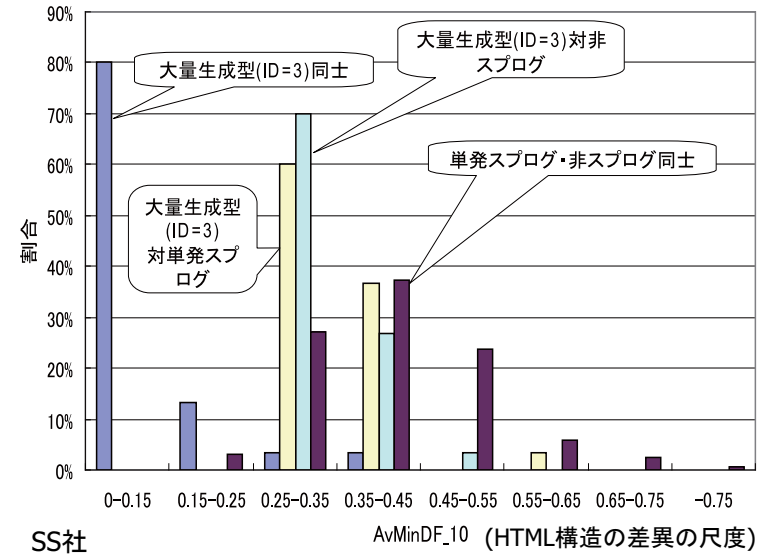
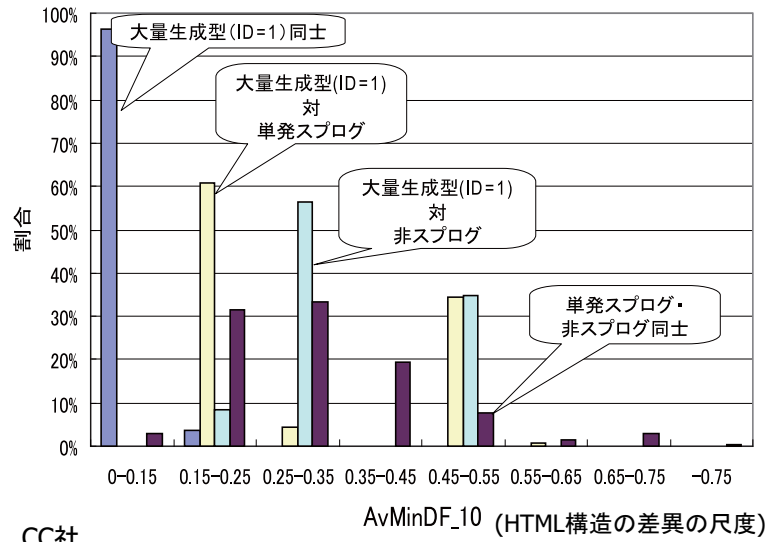


P78 機械学習を用いたスプログ検出における  
HTML構造の類似性の利用

片山 太一, 芳中 隆幸, 宇津呂 武仁,  
河田 容英, 福原 知宏

- HTML構造・・・HTML文書から抽出したDOM系列で表現
- 同一スパマーによって大量生成されたスプログ・・・DOM系列が類似
- DOM系列の有効性の検証
  - 教師あり学習によるスプログ検出
  - 教師なしスプログ収集

## DOM系列の類似



## SVMによるスプログ検出

