非文字データベースを対象としたAHPに基づくcovert channelの解析

電気電子情報工学科　4年B組
木下研究室　中村　恒太

背景

非文字データベースに着目する。

この時、資料(客体)へのアクセスを制御するために、主体とアクセスオペレーションによって記述するアクセス行列が用いられる。

従来、このようなアクセス制御のための資料の分類整理としてはオントロジーが使用されてきた。

問題

アクセス制御はアクセス行列により表現される。

アクセス行列には一般にcovert channelと呼ばれる情報漏洩を引き起こす経路が存在する。

しかし、資料と資料の関係を従来のオントロジーで表現する場合、covert channelの制御は困難を極める。

さらに、人と人との関係、あるいは人と情報との関係を論理学的モデルによって一般的かつ普遍的に示すことは、使い勝手を悪くするという問題があった。

目的

- Covert channelの分析において、従来のセキュリティモデルのようにオブジェクトを確率変数(単語)の集まりとみなし、トピック分析によるクラスタを分類項目とする。
- さらに、サブジェクト同士の関係には従来のセキュリティモデルの属性を使用する。
- 上記２つを組み合わせcovert channel解析する。

トピックモデル

- 文章が複数の潜在的なトピックから確率的に生成されると仮定したモデルであり、また、文章内の各単語はあるトピックが持つ確率分布に従って出現すると仮定します。
- トピックごとに単語の出現頻度を想定することで、トピック間の類似性やその意味を解析できる

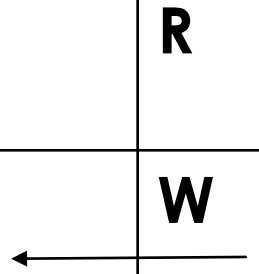
LDA(Latent Dirichlet Allocation)

文書中の単語の「トピック」を確率的に求める言語モデル
1つの文書が複数のトピックから確率的に構成されることを仮定した言語モデルの一手法です.

Covert channel

セキュリティポリシーに違反する抜け道。
・始点（Subject2・Object1）Subject2 が Object1 を読み込む
・中間点１（Subject2・Object2）Subject2 が Object2 に Object1 で読んだ内容を書き込む。
・中間点２（Subject1・Object2）Subject1 が Object2 を読む.
・終点（Object1・Subject1）Covert Channel により間接的にObject1 の内容を読めてしまう。

	S1	S2
O1	Φ	R
O2	R	W

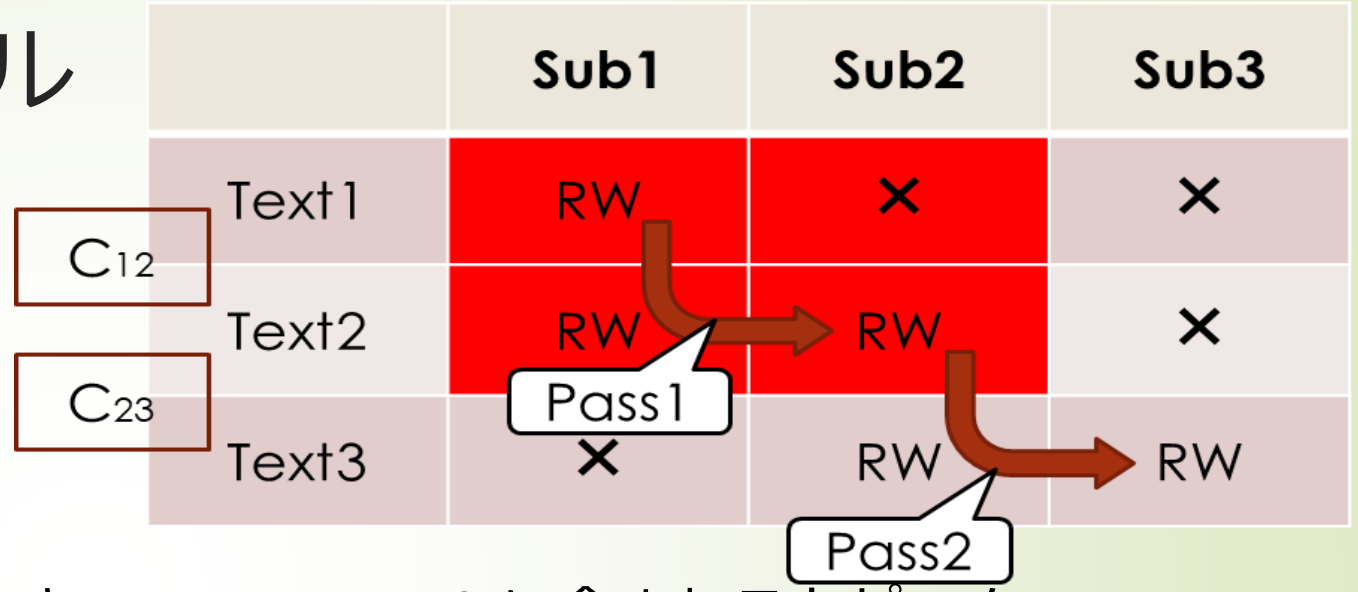


このように不正な情報流出が発生してしまうため、アクセス制御を行う推論エンジンとしては出来るだけ発生を抑制し、 検出と訂正を的確に行えるようにするのが情報フィルタに必要な機能である。

Coherence

- Coherence とはLDAにおいて、トピックの品質を測る評価指標である。
- LDAでは、文書に合計トピック数を設定すると自動でトピックを推測し、各文書の所属確率を計算できる。
- しかし、トピックモデルが推測したトピックの単語同士の一貫性は保証されない。そこでLDAではcoherence を算出し、トピック内に含まれる単語に一貫性があるかを計算する。
- coherence の算出にはトピック内の単語間の類似度が大きな役割を果たしている。よって、ここでは算出されたcoherence の値を単語間の類似度と定義する。

提案する概念モデル



1.covert channel の評価は subject 間の関係(役割、競合、所有) が関与すると同時にsubject 間を転送されるText 間の性質が考慮されなければならない。

2.Text 間の性質：Pass1に着目したとき、Text1、Text2 に含まれるトピック内の単語のcoherence によって判断する。これは、トピック分析から計算される。

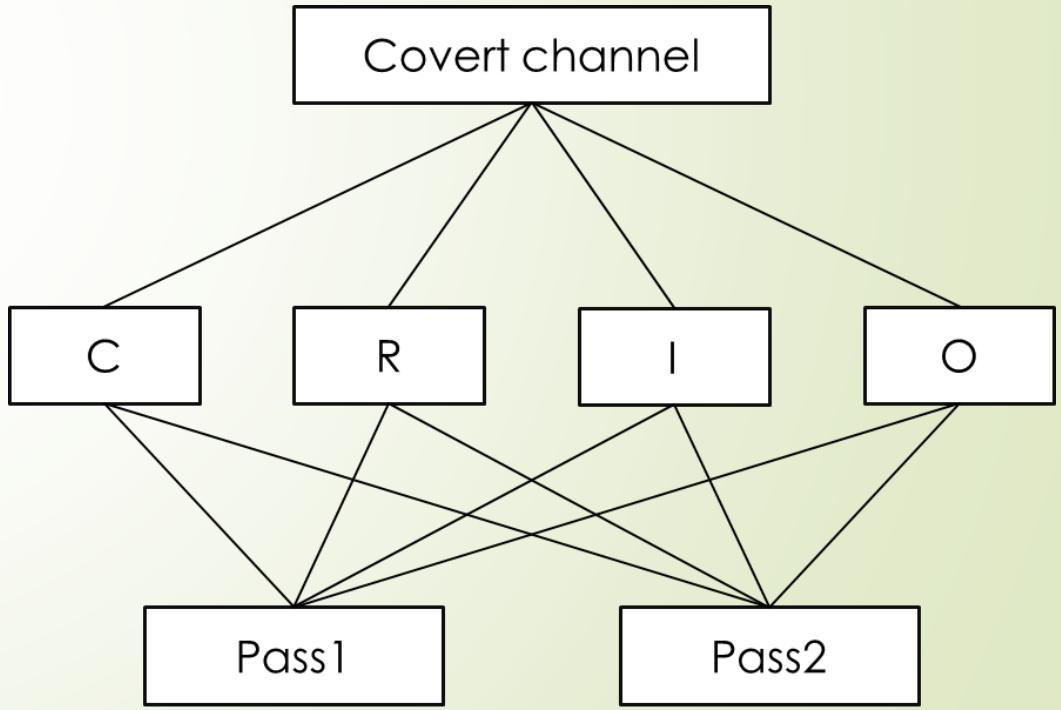
3.subject 間の関係及びText 間のトピックcoherence は主観的に評価されると前提する。

4.AHP の計算により、covert channel 上のフィルタリングすべき Pass の優先順位付けをする。

5.トピックモデルによる分析と情報セキュリティ属性(役割、競合、所有) がAHPで評価され、その結果からフィルタリングすべきcovert channel が決定される。

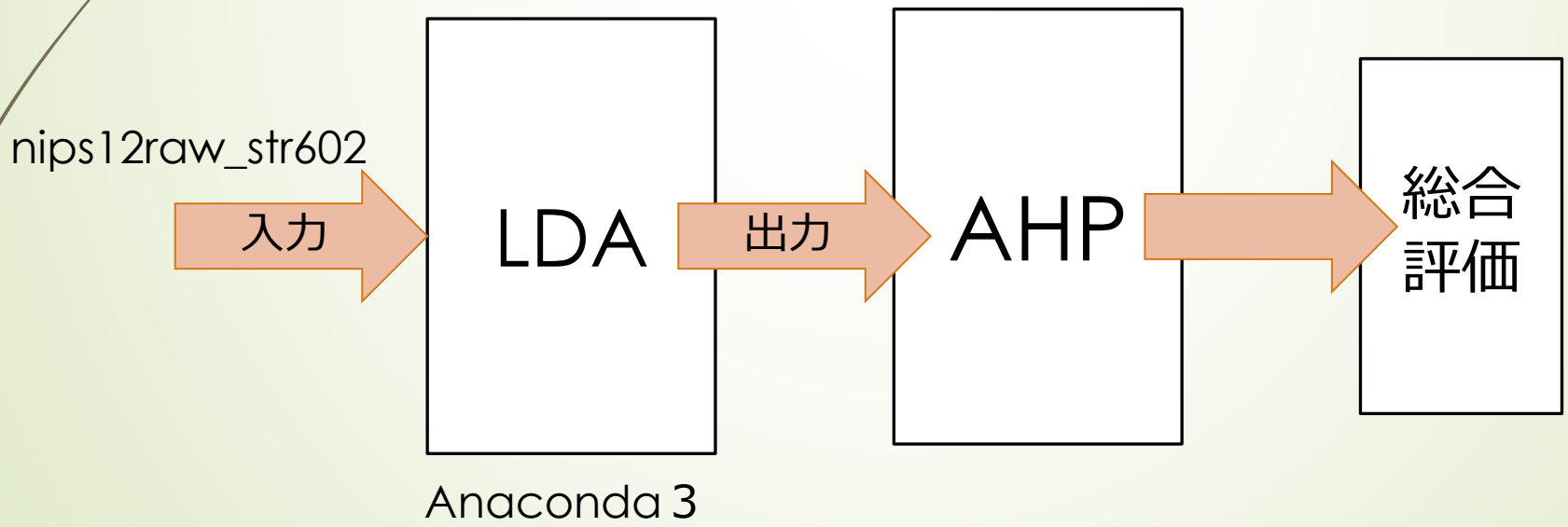
AHPによる優先順位付け

1．LDAにより、独立事象として定義されている単語の集合のトピック分布を計算する。
2．これにより計算されるcoherenceはトピック内の単語の類似度である。このcoherenceを使用し、AHP の評価項目(テクスト間の類似) とする。
3．さらにアクセス行列におけるcovert channel 上に出現するsubject 同士の関係を評価基準(R:役割、I:競合、O:所有) とする。
4．AHP の計算により、covert channel 上のフィルタリングすべきPass の優先順位付けをする。



実験

- Textとして論文データをLDAによりトピック分析し、トピックのcoherenceを出力する。
- AHPの重みを固有値法から算出するソフトで求める。
- その結果から総合評価を行う。



LDAによるトピック分析

- 大量のテクストをトピック分析し、単語20個ごとを10個のトピックに設定した出力結果の一部である。

全トピックのcoherenceの平均

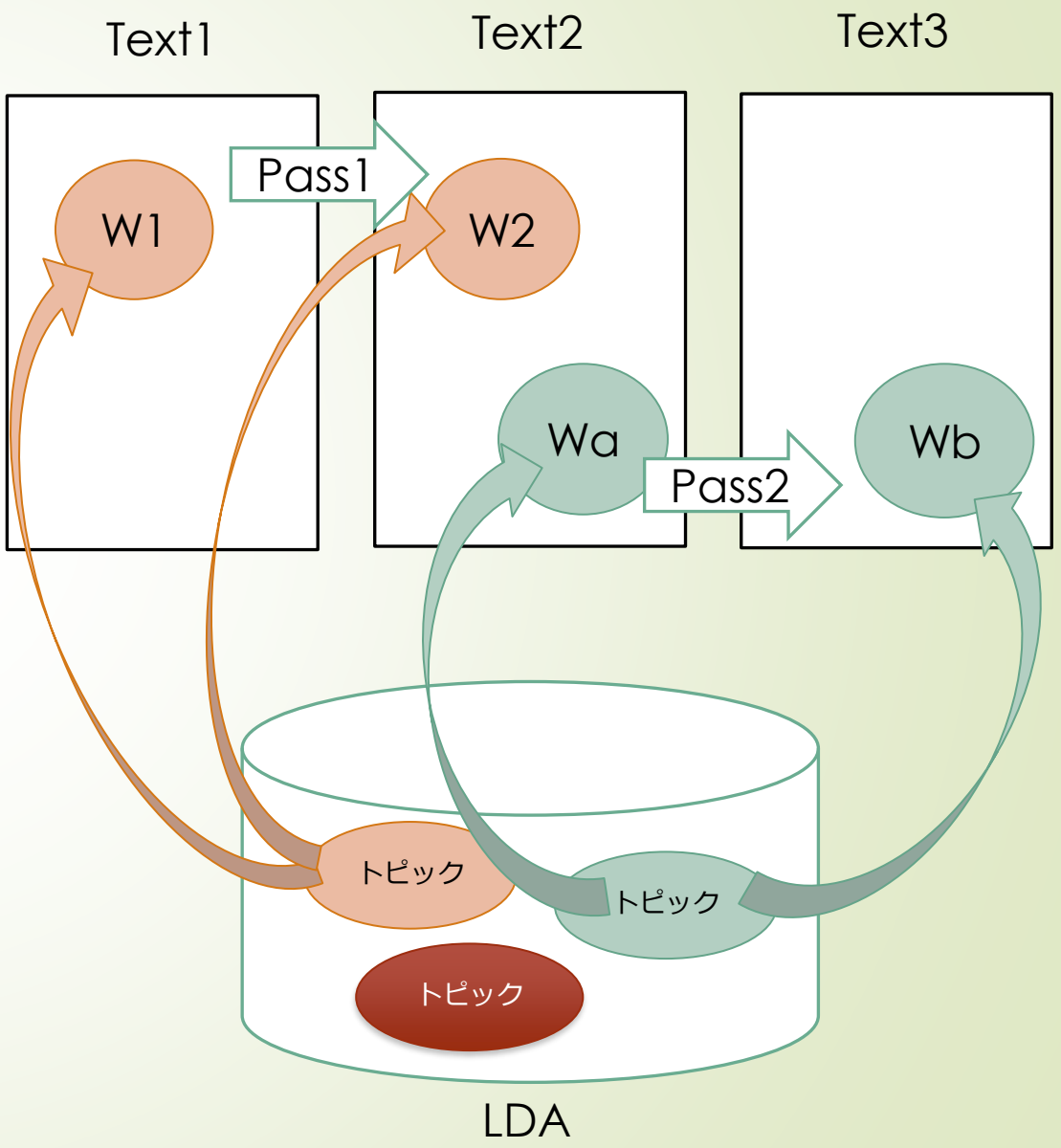
```
Average topic coherence: -1.1241.
[([(0.025163664, 'neuron'),
   (0.014695453, 'cell'),
   (0.009174355, 'spike'),
   (0.008574755, 'synaptic'),
   (0.007183699, 'firing'),
   (0.006625933, 'activity'),
   (0.005360948, 'connection'),
   (0.005293554, 'dynamic'),
   (0.004822483, 'response'),
   (0.004687287, 'potential'),
   (0.004228337, 'memory'),
   (0.003953116, 'synapsis'),
   (0.0038689172, 'fig'),
   (0.0038664965, 'simulation'),
   (0.0037337197, 'phase'),
   (0.0034825401, 'excitatory'),
   (0.0034173392, 'inhibitory'),
   (0.0032120293, 'signal'),
   (0.0031823071, 'membrane'),
   (0.0030939183, 'threshold')],
  -0.9630445183762313),
```

```
([(0.0052642035, 'net'),
  (0.005045612, 'hidden'),
  (0.0046278588, 'sequence'),
  (0.004625344, 'machine'),
  (0.004386533, 'solution'),
  (0.004208156, 'language'),
  (0.004180493, 'node'),
  (0.0038425317, 'string'),
  (0.0037875888, 'hidden_unit'),
  (0.0037045274, 'cost'),
  (0.003578985, 'optimization'),
  (0.00333463, 'constraint'),
  (0.0033199114, 'table'),
  (0.0033088576, 'recurrent'),
  (0.003233348, 'code'),
  (0.0031989065, 'symbol'),
  (0.003080977, 'activation'),
  (0.003000487, 'matrix'),
  (0.002989608, 'search'),
  (0.0026564174, 'grammar')],
 -1.4290562789759915)]
```

各トピックのcoherenceの値

AHP分析

- AHPの優先度の方針
- Pass1:coherenceの値が高く類似度が高いが、セキュリティモデルの属性はPass2と比べると優先度が低いものとする。
- Pass2:coherenceの値が低く類似度が低い。セキュリティモデルの属性はPass1よりも高いものとする。



AHP分析

- (1)セキュリティモデルの属性の重要度は個人によって変化するため自己で設定した。
- (2)評価基準の重要度を一対比較した結果から評価基準の重みを決定する。
- (3)評価基準ごとの代替案の一対比較を行い、評価基準ごとの代替案の評価を求める。

	C(類似度)	R(役割)	I(競合)	O(所有)	重み
C(類似度)	1	1/5	1/3	1/5	0.067935
R(役割)	5	1	3	1	0.389862
I(競合)	3	1/3	1	1/3	0.152352
O(所有)	5	1	3	1	0.389862

C(類似度)	Pass1	Pass2	score
Pass1	1	5	0.833333
Pass2	1/5	1	0.166667

I(競合)	Pass1	Pass2	score
Pass1	1	1	0.5
Pass2	1	1	0.5

R(役割)	Pass1	Pass2	score
Pass1	1	1/5	0.166667
Pass2	5	1	0.833333

O(所有)	Pass1	Pass2	score
Pass1	1	1/3	0.25
Pass2	3	1	0.75

AHP分析

(4)重みと評価基準ごとの代替案の評価から総合評価を出す。
Total scoreの高いほうが断ち切るべきPassとする。

	C(類似度)	R(役割)	I(競合)	O(所有)
重み	0.067935	0.389862	0.152352	0.389862

	C	R	I	O	total score
Pass1	0.833333	0.166667	0.5	0.25	0.2952311
Pass2	0.166667	0.833333	0.5	0.75	0.7047799

効果

- LDA(トピックモデル)とAHP(意思決定支援システム)の組み合わせにより論理学的なオントロジーによらず、テクストの確率測度的側面と人間の主観を組み合わせたcovert channel解析のためのモデルが提案された。

- これにより、個人または集団にマッチした情報漏洩の制御が可能になる。