情報漏洩防止のための
アクセス制御システムの設計

木下研究室　201503917　松内滉征

研究背景

・近年、SNSやオンラインストレージなどのサービスの普及により情報を簡単に便利に扱えるようになっている。

・情報収集者は本来は制限されているはずの情報を様々な情報リソースを統合し、推論することにより取得してしまう可能性がある。

問題点

・インターネットにおいて許可されたオブジェクトのみをアクセス可能にしたにもかかわらず推論により制限されたオブジェクトの内容が漏洩してしまう可能性がある。

・検索時の推論による情報漏洩の問題に着目する。

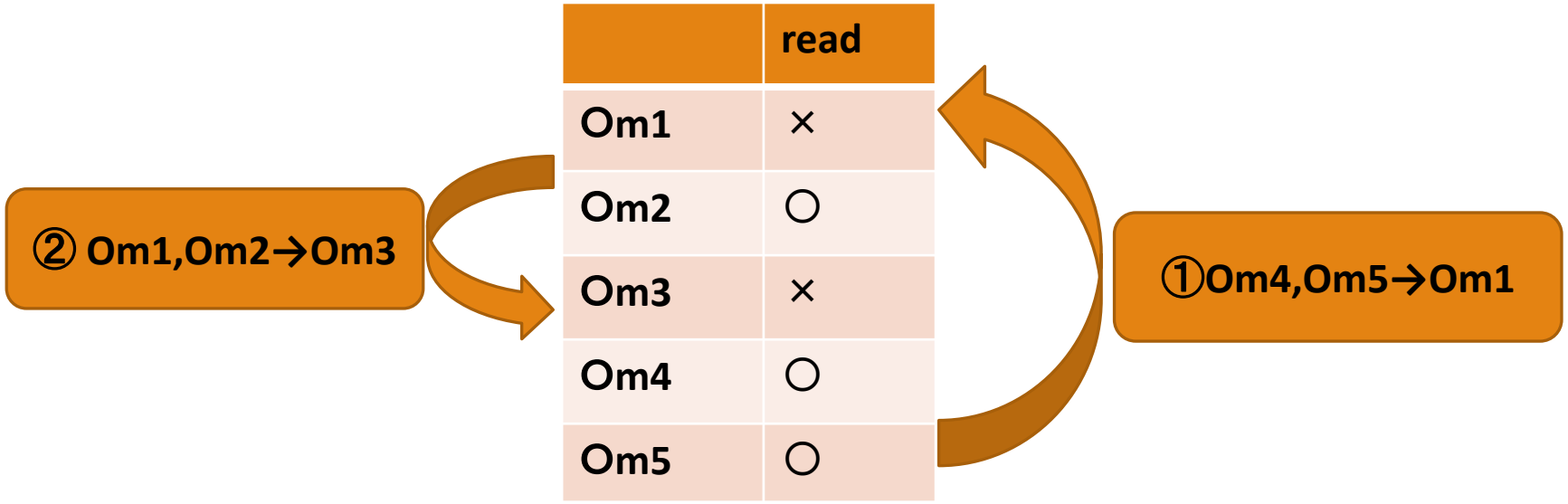
目的

全文検索とストレージの間に推論攻撃を防止する情報フィルタを設計し、ユーザごとの推論攻撃に対応できるアクセス制御システ ムの提案をする。

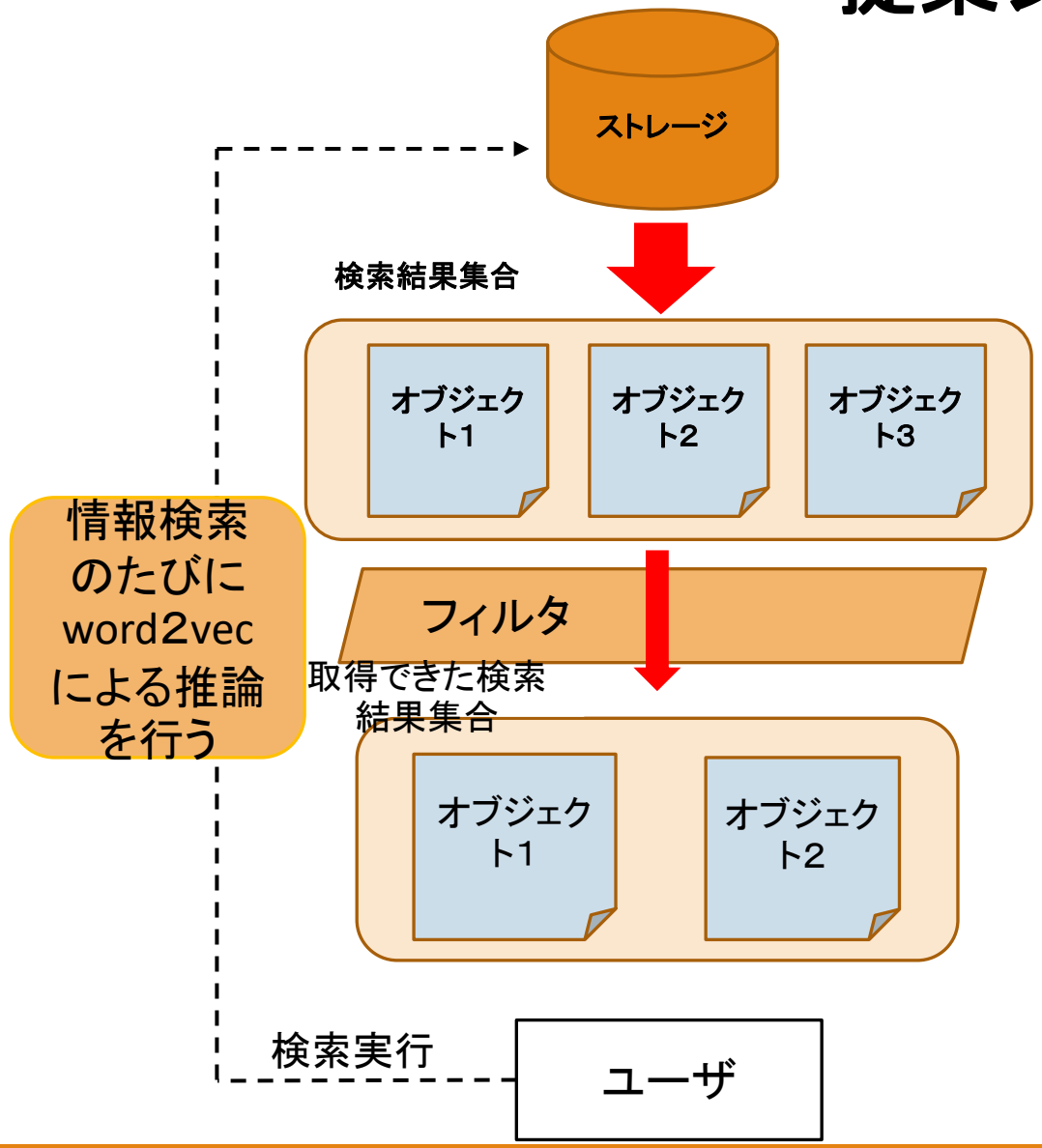
推論による情報漏洩

読み取り可能なオブジェクトOm2,Om4,Om5読み取り不可のオブジェクトOm1,Om3が存在する。

推論により図の①∧②が可能であるとき本来読み取ることができなかった二つのオブジェクトの両方が読み取ることができてしまい情報漏洩となる。



提案システムの概念図



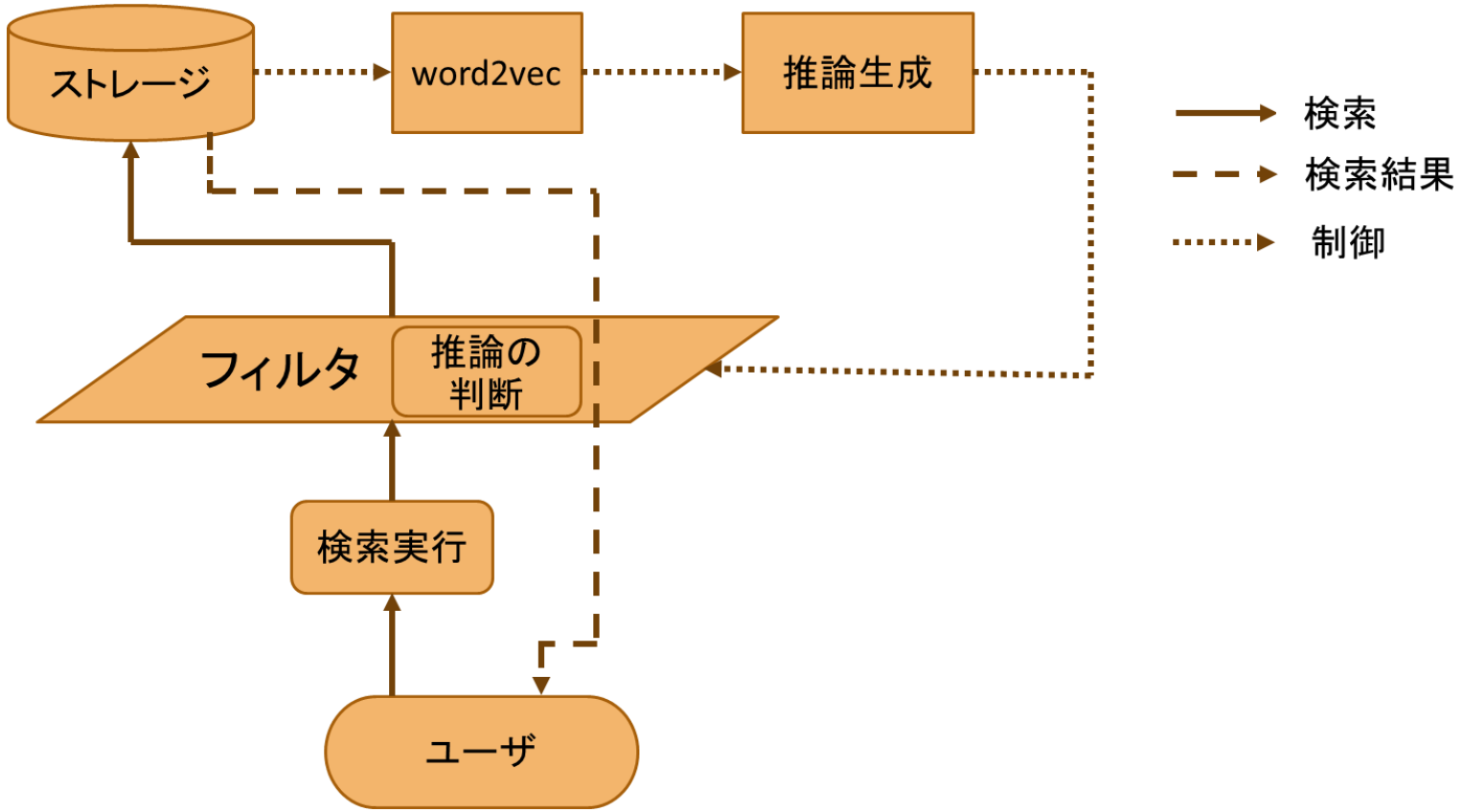
推論により情報漏洩につながるオブジェクトをブロックするようにフィルタの更新を行う。

オブジェクト3の情報を推論させないためにオブジェクト1、オブジェクト2または両方をブロックする。

提案システム構成要素

- **OS**：ubuntu18.04.1
- **使用言語**：python 3.6.6 (anaconda 1.6.9)
- **whoosh**：Python純正の全文検索エンジンのライ ブラリ。
- **ストレージ**：データが保管されている場所。Ubuntu のディレクトリ。
- **word2vec**：単語をベクトル表現化し、単語の関係性をベクトル演算により表現できる。
- **使用テキスト**：Wikipedia　2.3GB（MeCabで分かち書き）
- **推論規則生成**：word2vec の結果を利用した推論規則を生成する。
- **フィルタ**：推論規則に基づきwhooshの検索結果を制御する。

推論予測フィルタによる検索システムの動作



システムの流れ

1.ユーザがwhooshで全文検索する。

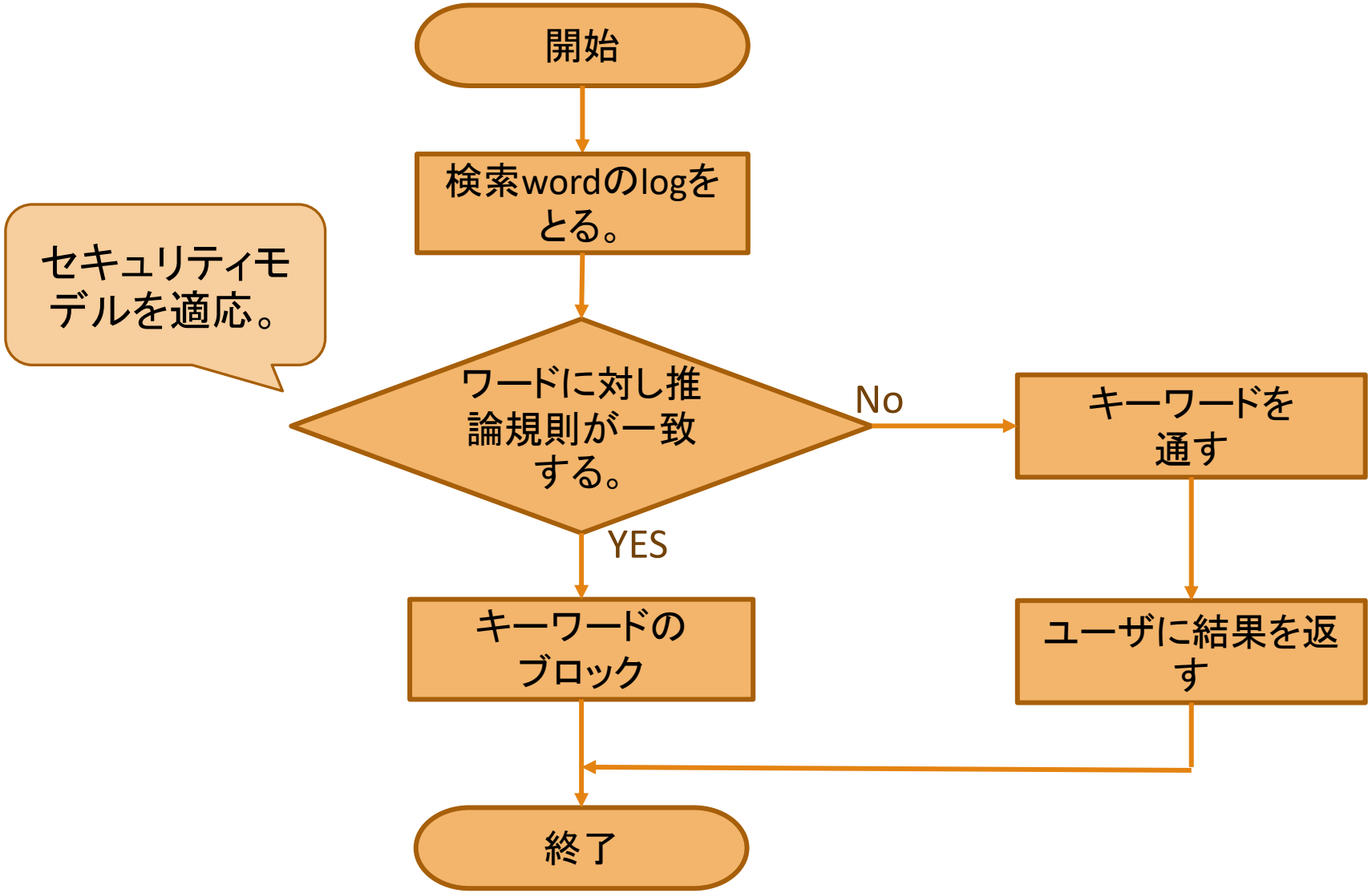
2. 検索により文章の中に指定の単語が含まれているか判断する。

3. 機密情報に近い単語をword2vecの推論規則により判定する。

4. フィルタにより情報漏洩に繋がる恐れがあるデータのアクセス制御を行う。

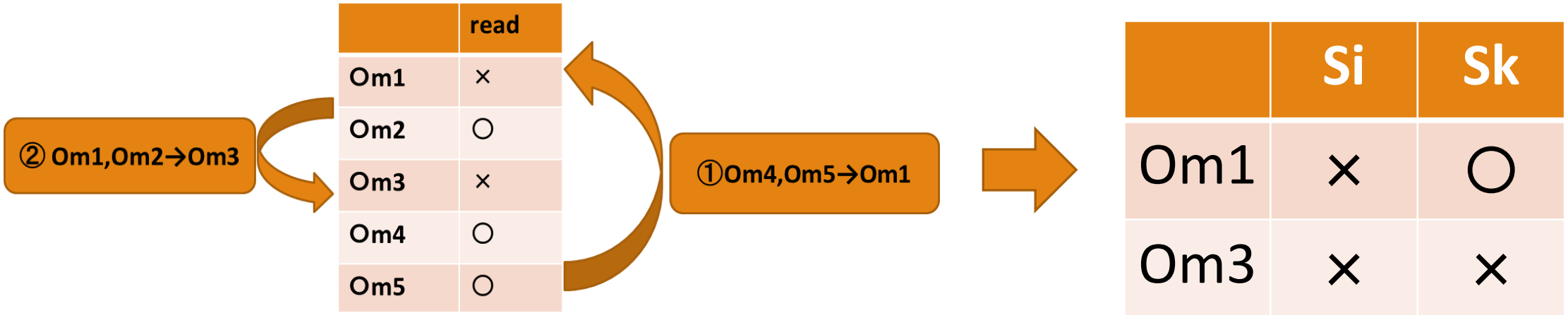
5. 1～4をユーザの検索のたびに行う。

検索による推論防止フィルタ
機能フローチャート



情報フィルタのためのセキュリティモデル

読み取り不可の情報の中でもユーザによってその情報の価値は異なる。
そのためブロックする情報もユーザごとに適応させる必要がある。
読み取り不可であるOm1,Om3の情報が検索するユーザにより機密情報になりえない場合を想定する。



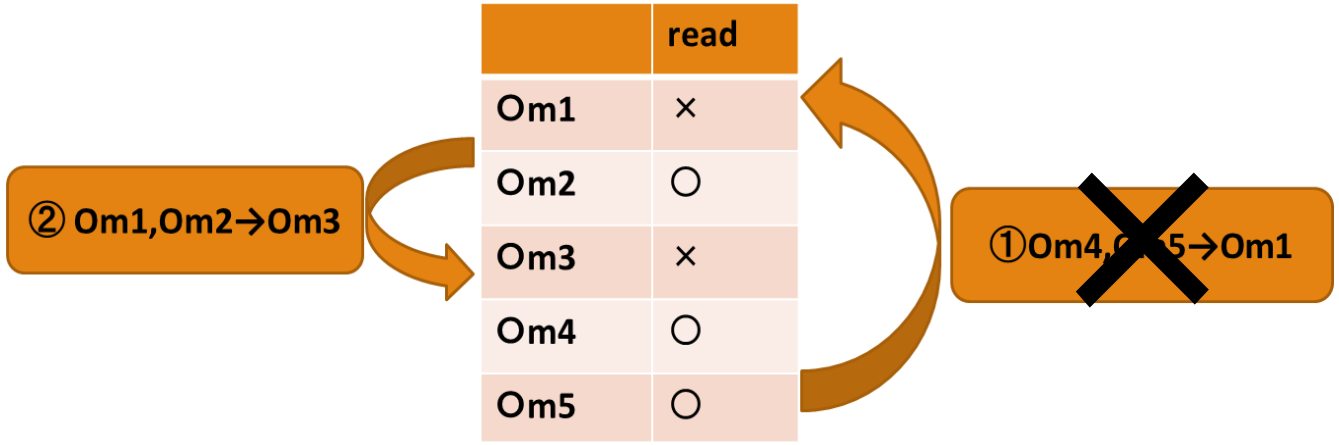
	read
Om1	×
Om2	○
Om3	×
Om4	○
Om5	○

	Si	Sk
Om1	×	○
Om3	×	×

情報フィルタのためのセキュリティモデル

以下のような操作をすればこのモデルにおける推論攻撃を防止できる。

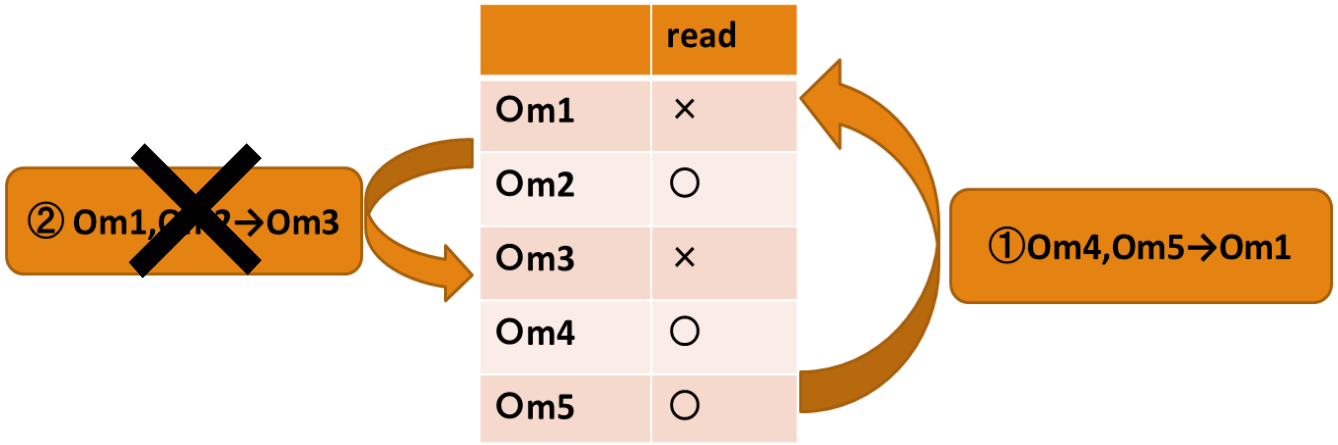
（1）S1に対して
図の①の推論を防ぐ。Om1の情報がなければ②の推論も不可能である。
そのためOm4∨Om5∨{Om4∧Om5}をブロックする必要がある。
なぜならSiにとってOm1,Om3どちらの情報も機密情報になりえるからである。



	Si	Sk
Om1	×	○
Om3	×	×

情報フィルタのためのセキュリティモデル

（2）Skに対して
図②の推論を防ぐ。
そのためOm2のブロックをする必要がある。
Om2の読み取りを不可にすることにより、Om4,Om5の情報は読み取り可とし①の推論によりOm1の情報は読み取ることが可能になった場合にも、Skにとっての機密情報であるOm3を推論により読み取ることを防ぐことができる。



	read
Om1	×
Om2	○
Om3	×
Om4	○
Om5	○

	Si	Sk
Om1	×	○
Om3	×	×

結果･考察

提案システムの要素であるwhooshとword2vecを実装し、機能を確認した。

セキュリティモデルを考慮したフィルタを設計した。

今後の課題はこれにより情報漏洩を防ぐ検索システムで正常にアクセス制御できるかどうか検証することである。