

SNSにおける 関係の意味情報モデルの提案 ～評価方法について～

南山大学 数理情報学部 情報通信学科

2008MI049 久田晋也

2008MI054 細川 翼

指導教員: 青山幹雄

- ◆ 関係の意味的距離
- ◆ 関係の意味的距離の問題点
- ◆ 関係の意味的距離の再定義
- ◆ 類似性に基づく友人推薦手法
- ◆ 評価方法
- ◆ 今後の課題

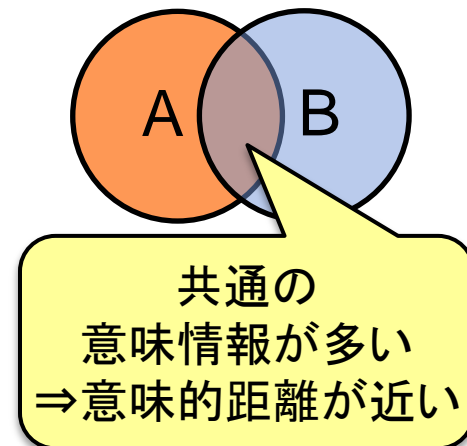
◆中間発表で提案した定義

A: 友人紹介サービス利用者の持つ意味情報集合

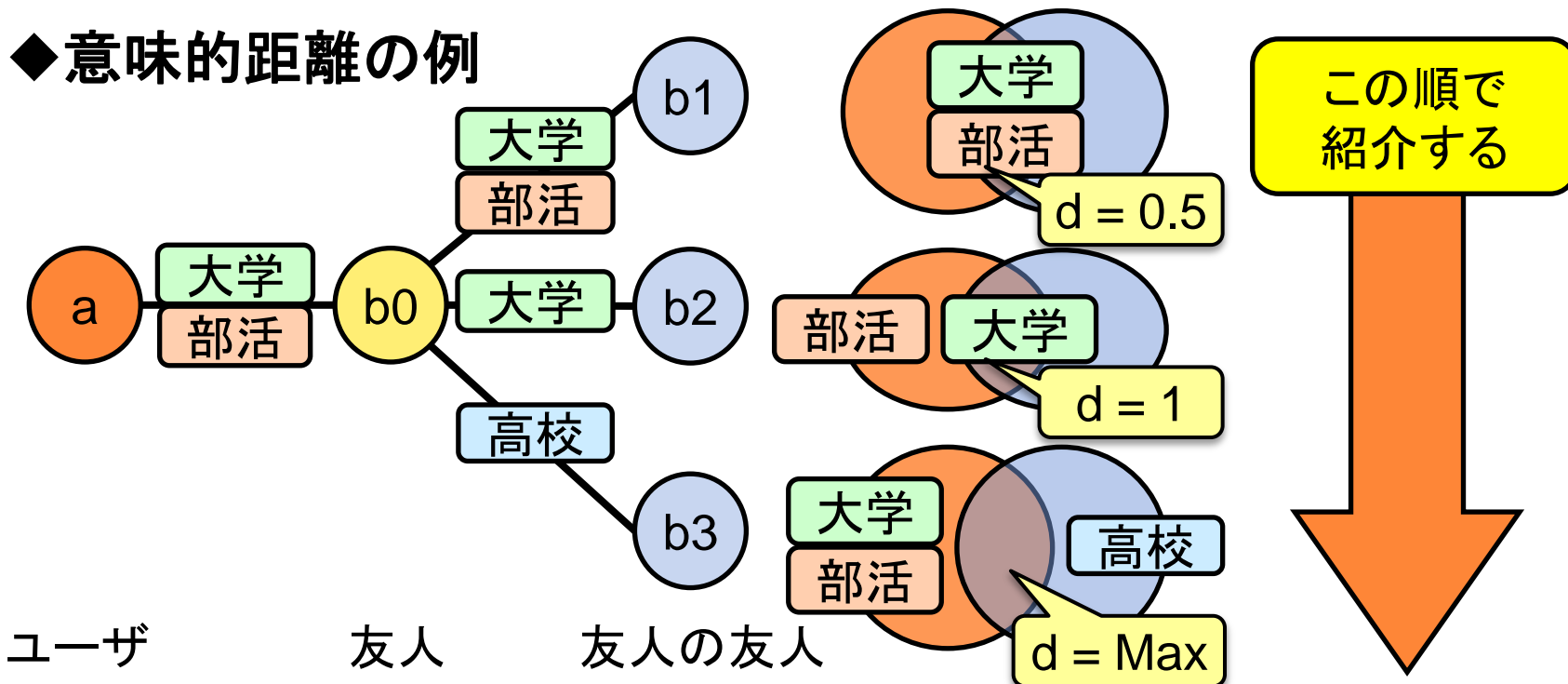
B: 紹介される友人の持つ意味情報集合

$$d = \frac{1}{|A \cap B|}$$

ただし、 $|A \cap B| = 0$ の場合は $d = \text{Max}$ とする

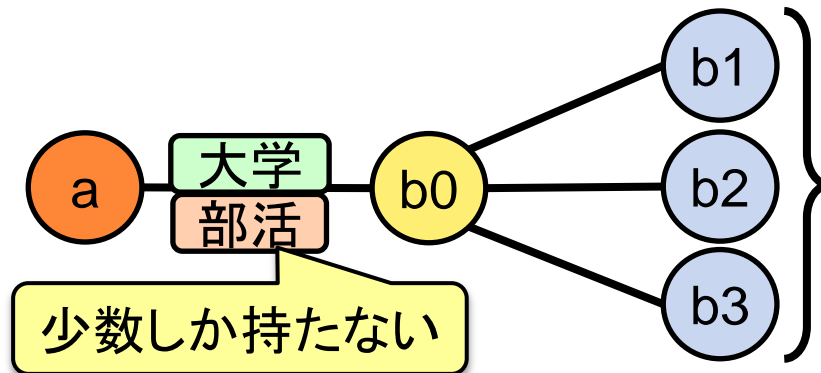


◆意味的距離の例



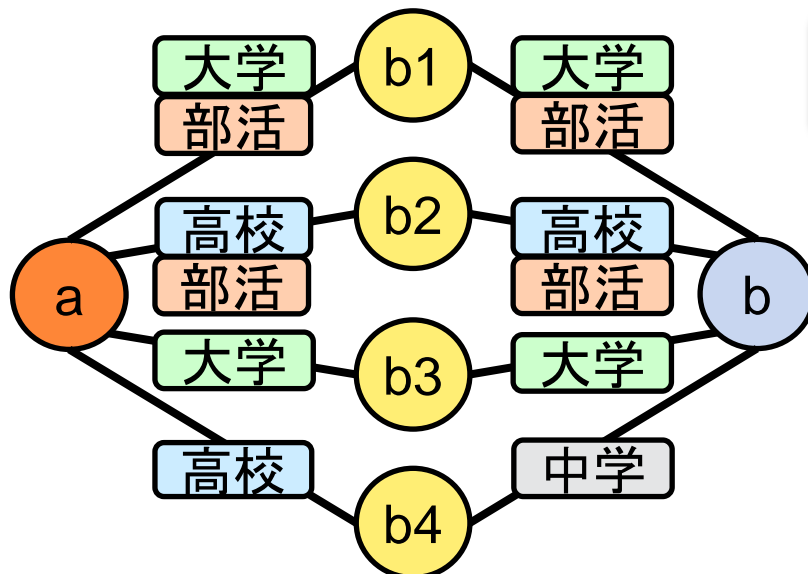
◆問題点

✓ 実際の関係を考えて、関係の意味情報は少数通りでしか考えられない



$|A \cap B|$ は0, 1, 2
しか考えられないので
 d は3通りの表現しかできない

✓ 共通の友人が複数いる場合が考慮されていない



誰を共通の友人として計算するかで
 d の値が変化する

共通の友人	$ A \cap B $	d
b1	2	0.5
b2	2	0.5
b3	1	1
b4	0	Max

関係の意味的距離の再定義

◆関係の意味情報の再定義

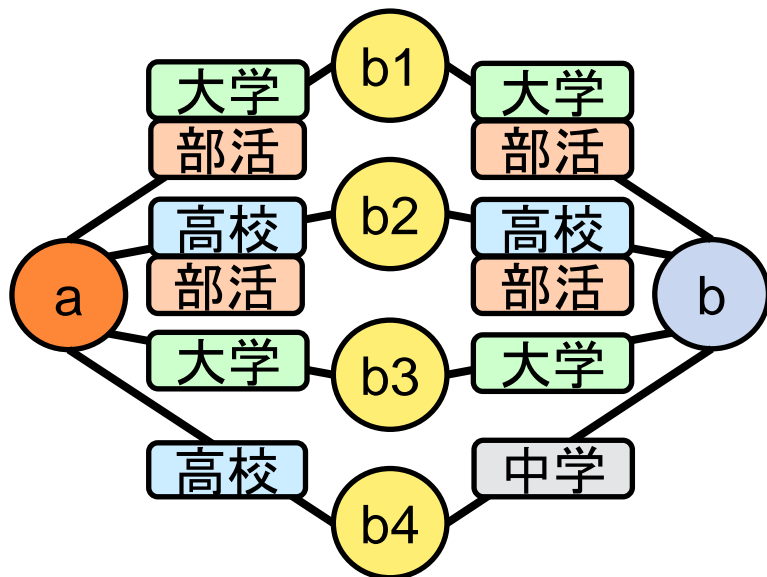
$$\vec{a} = (a_1, a_2, \dots, a_n)$$

(n : 共通の友人数, a_j : 共通の意味情報数)

とすると、

$$d = \frac{1}{|\vec{a}|} = \frac{1}{\sqrt{a_1^2 + a_2^2 + \dots + a_n^2}}$$

ただし、 $|\vec{a}| = 0$ のとき、 $d = \text{Max}$



共通の友人	$ A \cap B $
b1	2
b2	2
b3	1
b4	0

$$\vec{a} = (2, 2, 1, 0)$$

$$d = \frac{1}{3}$$

◆2つの推薦手法の提案と評価

◆1. 個人内情報を利用

- ✓合成友人プロフィールベクトルAPFと候補ユーザの
プロフィールベクトルANF_jの余弦の大きいユーザから紹介
- ✓ベクトルの要素はある属性の要素値に対応
例 ギターが趣味で読書は趣味でないユーザ (ギター, 読書)=(1, 0)

◆2. 個人間情報を利用

- ✓合成友人コメントベクトルCPFと候補ユーザのコメントベクトルCNF_j
の余弦の大きいユーザから紹介
- ✓ベクトルの要素はコメント数とコメントしたユーザに対応
例 Bに2回, Cに1回コメントしたユーザ (B, C)=(2, 1)

◆評価に使用したデータ

- ✓イントラSNSのソーシャルグラフを利用
- ✓OpenPNEで構築
 - OpenPNE: オープンソースのSNS運営用サーバサイドプログラム群
- ✓2010年7月運用開始, 2010年12月のデータで分析

◆評価

- ✓現実で友人・SNSで友人でないユーザが紹介されるかどうか
- ✓被験者に自分の友人を列挙してもらうのは, 倫理的に問題
- ✓同じ研究室所属ユーザは友人と仮定
- ✓推薦順位の上位5人に同じ研究室ユーザが何人いるかで評価

◆考察

- ✓提案手法の有効性を厳密に評価することはできない
- ✓利用者は全体の6割

◆ 関連研究の方法を使用すると…

- ✓ 瀬戸キャンパスのSNSの構築, 試験運用, 情報取得
⇒ 期間が必要
- ✓ 友人の仮定
 - 研究室内としたとき, 高校の友人・趣味などの属性は紹介に相関しないと考えられる
 - 『XX研究室所属』は関係の意味情報に付加すべき情報

◆ 試験データ

- ✓ 架空のソーシャルグラフを作成することを検討
 - 実際のSNSはプライバシーなどの問題が発生, 困難と考えられる
 - 架空のソーシャルグラフに対し提案方法を適用, 従来の方法と比較評価

◆関係の意味情報の検討

- ✓属性の分類化や属性の重みなど詳細検討
 - 友人より家族の重みを大きくするなど
- ✓意味的距離も再検討の余地がある
 - 余弦使用の検討

◆試験環境の検討

- ✓架空ソーシャルグラフの作成基準
 - ユーザ数, エッジの属性数と意味情報など

SNSにおける 関係の意味情報モデルの提案 ～評価方法について～

END

南山大学 数理情報学部 情報通信学科

2008MI049 久田晋也

2008MI054 細川 翼

指導教員: 青山幹雄