

SNSのソーシャル性がICNのキャッシュ性能に与える影響の分析

Investigating Social Effect of SNS on Cache Performance of ICN

上山憲昭¹ 本田紘大² 中村遼³

Noriaki Kamiyama Kodai Honda Ryo Nakamura

立命館大学 情報理工学部¹

College of information Science and Engineering, Ritsumeikan University

立命館大学大学院 情報理工学研究科²

Graduate School of Information Science and Engineering, Ritsumeikan University

福岡大学 工学部³

Faculty of Engineering, Fukuoka University

1 まえがき

今日では、Twitter (X) や Facebook に代表されるソーシャルメディア (SNS: social networking service) は、我々の社会生活を支える不可欠な基盤となっている。SNS を介して我々はオンライン上でも友人と交友を深められ、社会活動や経済活動をより豊かにしている。SNS では動画データも頻繁に交換されるため、SNS の普及に伴い通信ネットワーク上のトラフィック量が增大することから、コンテンツ転送の手段として情報指向ネットワーク (ICN: information-centric networking) が有望視されている [1]。ICN ではルータはコンテンツの複製を一時的に保持でき、中継ルータからもコンテンツが配信可能であるため、ネットワーク上を転送されるトラフィック量が削減される。ただし ICN を SNS のコンテンツ配信基盤とする上では、SNS に参加するユーザーの特性 (ソーシャル性) を考慮した ICN の設計・制御が望ましい [2]。ICN で重要なのは、ルータのキャッシュへのコンテンツ注入判断である「キャッシュ挿入」および「キャッシュ置換」を適切に設計することである。これらのキャッシュ制御にユーザーのソーシャル性を考慮することができれば、ソーシャル性を指向した ICN の実現につながる。

ユーザーのソーシャル性の主要なものは以下のようにまとめられる [2]。

- **インフルエンサの存在:** 影響力の強いユーザーが発信した情報は SNS 上で拡散されやすい
- **情報拡散行動:** あるユーザーが発信した情報が、ユーザー同士の交友関係 (ソーシャルネットワーク) を介して他のユーザーに伝搬
- **コミュニティ:** 嗜好や性質が似たユーザーらによりコミュニティが形成され、同一コミュニティ内の交友は活発

そこで我々はこれまでに、インフルエンサの作成したコンテンツのみをルータでキャッシュする場合 [3] に、インフルエンサの選択方法やその割合がキャッシュヒット率に与える影響を分析した [4]。またユーザーの情報拡散行動やコミュニティ構造がキャッシュヒット率に与える影響を分析した [5]。本稿では、これら分析の概要をまとめる。

2 ソーシャル性を考慮した ICN モデル

2.1 ネットワークモデル

キャッシュを有する ICN ルータから構成されるネットワークを無向グラフ $G_I = (V, E)$ で表記する。ICN ルータは要求およびコンテンツを中継するとともに、キャッシュ挿入方式とキャッシュ置換方式に従ってコンテンツをキャッシュする。また SNS 上のユーザー間の交友関係を表すネットワーク (ソーシャルネットワーク) を有向グラフ $G_S = (U, L)$ で表記する。有向グラフ G_S 上における辺 (u, v) は、ユーザー u がユーザー v をフォローしていることを意味する。ユーザー u がフォローしているユーザーの集合 \mathcal{N}_u^+ は $\{v | (u, v), v \in U\}$ であり、ユーザー u をフォローしているユーザー (フォロワー) の集合 \mathcal{N}_u^- は $\{v | (v, u), v \in U\}$ である。

ユーザーが属するコミュニティの集合を C と表記し、ユーザーはいずれか 1 つのコミュニティ $c \in C$ に属するものとする。SNS 上のユーザー $u \in U$ の収容先となる ICN ルータ $r \in V$ をユーザーが属するコミュニティに基づいて決定する。まずコミュニティ単位で収容される ICN ルータをランダムに決定し、そのコミュニティに属する全てのユーザーを収容する。その後、各ユーザーを一定の確率 (転移確率) でランダムに選択された異なる ICN ルータに転移させる。

2.2 ユーザー行動モデル

SNS 上のユーザーの行動モデルとして、[6] で提案されている行動モデルに、ユーザーの情報拡散行動を導入したものをを用いる。ユーザー行動モデル [6] ではユーザーは以下のように振る舞う。(i) セッション開始、(ii) セッション内において、フォローしているユーザーが発行したコンテンツを閲覧、もしくは自身が新たにコンテンツを発行、(iii) 一定の間隔を空けながら (ii) を反復後にセッション終了、(iv) (i)~(iii) を一定の間隔を空けながらセッション単位で反復。

本稿では (ii) における行動を以下のように拡張する。

- **要求:** ユーザー u は自身がフォローしているユーザー $v \in \mathcal{N}_u^+$ が最後に発行したコンテンツと、 v が広告したコンテンツを要求
- **広告:** ユーザー u はコンテンツを取得後に、自身をフォローしているユーザー \mathcal{N}_u^- に、確率 p_u でコンテンツを広告
- **発行:** ユーザー u は新たにコンテンツを発行

3 性能評価

3.1 インフルエンサの存在の影響分析

3.1.1 解析評価

SNS 上のユーザー間の交友関係のグラフ G_S の次数分布 $P(k)$ を与えたときの、コンテンツ要求のうち要求コンテンツが経路上のルータのキャッシュから返送された割合であるキャッシュヒット率の平均値を解析的に導出した [4]。ただしユーザーは各 ICN ルータにランダムに収容され (転移確率が 0)、ルータのキャッシュ挿入はインフルエンサの発行コンテンツのみをキャッシュする SACS [3] を用い、キャッシュ置換法は LRU を用い、キャッシュサイズは 100 [content] とする。ICN ネットワークにおいて、要求およびコンテンツが転送されるルータ間の経路は最短経路とする。

SNS の次数分布 $P(k)$ をポアソン分布、指数分布、べき分布とした場合の平均キャッシュヒット率をインフルエンサの割合 p に対して図 1 にプロットする。 $P(k)$ がポアソン分布や指数分布の場合は、キャッシュヒット率のピーク値と $p = 1$ の場合の他の場合で差がほとんどない。つまりこれらの次数分布の場合には、影響力の強いユーザーを決定することによるキャッシングへの恩恵は限定的である。一方、次数分布がべき分布である場合には、影響力の強いユーザーの割合 p に対してキャッシュヒッ

ト率は $p = 0.02$ 付近において上に凸となる。従ってインフルエンサの割合を適切に決定することで、 $p = 1$ の SACS に相当する従来のキャッシュ制御方式と比較して、キャッシュヒット率の大幅な改善が期待できる。

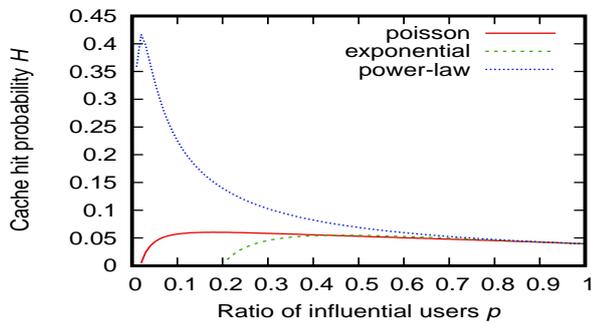


図1 SNSの次数分布による平均キャッシュヒット率の比較

3.1.2 シミュレーション評価

ルータ間の接続関係を表すグラフとして ER (Erdős-Rényi) モデルにより生成したランダムグラフ (ノード数 100, リンク数 200) を、ユーザ間の交友関係を表すグラフとして Last.fm (ノード数 1,843, リンク数 12,268) および Facebook (ノード数 4,039, リンク数 88,234) のデータセットを用いる。ルータのキャッシュ挿入は SACS を用い、キャッシュ置換法は LRU を用いる。またキャッシュサイズは 10 [content] とする。ICN ネットワークにおいて、要求およびコンテンツが転送されるルータ間の経路は最短経路とする。また転移確率は 0 (コミュニティ構造が無い) に、拡散確率は 0 に設定する。

与えた中心性指標により各ユーザの中心性を求め、中心性の高い上位 p 割 ($p = 0.1$) のユーザをインフルエンサとする。8 種類の中心性指標 (次数中心性、媒介中心性、近接中心性、固有ベクトル中心性、PageRank, k コア指標, VoteRank, CI (Collective Influence)) を用いる。

8 種類の中心性指標を用いたときの、Last.fm における平均キャッシュヒット率を図 2 に示す。図中には比較のため、ランダムに影響力の強いユーザを選択したときの結果もあわせて示す。中心性指標に基づいて影響力の強いユーザを選択することで、キャッシュヒット率が大幅に改善される。キャッシュヒット率は、特に媒介中心性、PageRank, VoteRank を用いた場合が最も高い反面、固有ベクトル中心性や k コア指標を用いた場合が最も低い。Facebook の場合も、同様の結果が得られた。媒介中心性や PageRank を求めるためには、ネットワークの規模に応じた計算量が必要となるため、大規模な SNS を対象とする場合には、計算量が少ない次数中心性が有効である。

3.2 情報拡散行動とコミュニティの影響分析

ICN ネットワーク G_I は、やはり ER モデルによって生成したランダムグラフを用いる。ただしノード数は 25, 平均次数は 3 である。SNS 上のソーシャルネットワーク G_S は、スケールフリーネットワークの生成モデルである Li-Maini モデルにより生成する。ノード数 (ユーザ数) は 1,000, 平均次数は 12.21 である。ICN ルータのキャッシュサイズは 5 [content] とし、キャッシュ挿入法は LCE (Leave Copy Everywhere) を置換法は LRU を用いる。

図 3 に転移確率 (コミュニティの強度) を変化させたときの、キャッシュヒット率を、拡散確率の 3 つの値の

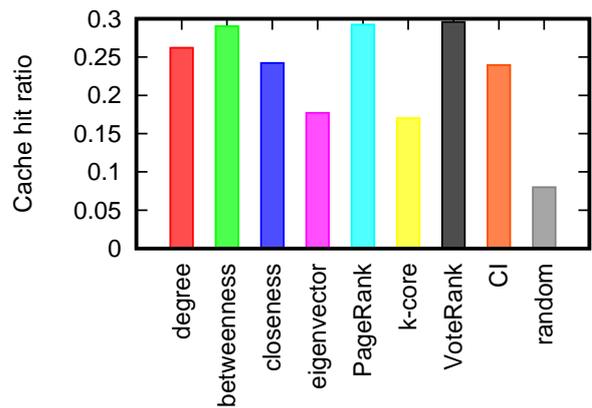


図2 各中心性指標の平均キャッシュヒット率 (Last.fm)

場合について各々示す。転移確率の増加 (コミュニティ構造の減少) に伴いヒット率は低下する。また拡散確率の増加に伴いキャッシュに挿入されるコンテンツの種類が増加するため、ヒット率は低下する。拡散確率の増加によるヒット率の低下は、コミュニティが保持されている方が顕著となる。

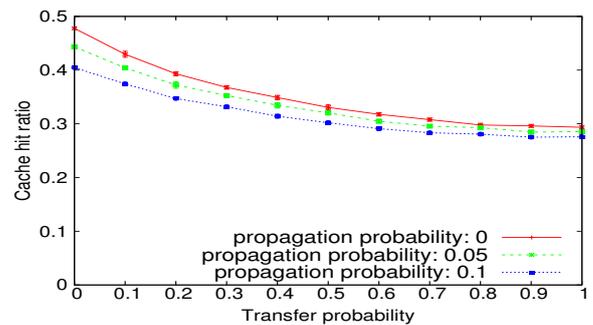


図3 コミュニティの強度に対するヒット率

謝辞 本研究成果は JSPS 科研費 21H03436 と 21H03437 の助成を受けたものである。ここに記して謝意を表す。

参考文献

- [1] G. Xylomenos, et al., A survey of information-centric networking research, IEEE Communications Surveys & Tutorials, vol. 16, no. 2, pp. 1024–1049, 2014.
- [2] S. Tsugawa, A survey of social network analysis techniques and their applications to socially aware networking, IEICE Transactions on Communications, vol. E102–B, pp. 17–39, Jan. 2019.
- [3] C. Bernardini, et al., Socially-aware caching strategy for content centric networking, IFIP Networking 2014.
- [4] K. Honda, et al., Analyzing Effects of Social Media User's Influence on Contents Caching in ICN, IEEE Access, Vol. 11, pp. 127679–127688, Nov. 2023.
- [5] 本田 紘大, 他, ユーザのソーシャル性が ICN のコンテンツキャッシングに及ぼす影響 -情報拡散行動とコミュニティに着目して-, 信学会 CQ 研究会, CQ2023-48, 2023 年 11 月.
- [6] F. Benevenuto, et al., Characterizing user behavior in online social networks, ACM IMC 2009.