

# Sharing Content at the Edge of the Network Using Game Theoretic Centrality

Anubhab Banerjee<sup>1</sup>, Nishanth Sastry<sup>2</sup>, Carmen Mas Machuca<sup>1</sup>

<sup>1</sup>Technical University of Munich (TUM), Germany

<sup>2</sup>Dept. of Informatics, Kings College London (KCL), UK

e-mail: anubhab.banerjee@tum.de

## ABSTRACT

Content Delivery Networks aim at delivering the desired content to each user at minimum delay and cost. To tackle this problem, the content placement problem considering available cache locations has been widely studied. However, this paper addresses this problem by taking advantage of using existing but still underused Wi-Fi links. Our study considers to cache content in user homes and sharing it among neighbours via Wi-Fi links. To maximize energy savings and reduce delays, content should be intelligently placed at the caches distributed in different users' homes. We propose using a 'game theoretic centrality' metric, which models the sharing of content among neighbours as a co-operative coalition game. We apply this metric to study the energy savings and evaluate how close the contents are placed to the interested user(s).

**Keywords:** Centrality measures, Content sharing, Edge caching, Game theoretic centrality

## 1. INTRODUCTION

Designs for future networks such as 5G, calls for more capacity or completely new technologies at the edge of the network such as mmWave [1]. We seek a different, more sustainable approach that makes use of *underused existing* resources more effectively. In particular, current versions of Wi-Fi such as 802.11ac and 802.11n offer 450 Mbps within a close space. These speeds exceed by far the commonly available broadband connection speeds in developed countries today<sup>1</sup> suggesting that it is possible to efficiently distribute content by *sharing content at the edge*. Based on this idea, Raman *et al.* [2] have recently introduced the *Wi-Stitch* architecture, showing the feasibility of sharing content at the edge of the network. Wi-Stitch proposes to stitch together a Content Delivery Network (CDN) by linking together users' homes using Wi-Fi (or similar technology), creating a local cell of users who can share content among each other. This sharing is enabled by provisioning storage for the shared content at users' homes. We envision that the cache would be attached to e.g., the ISPs modem in the users home, or a media streaming device operated by the content provider (e.g., versions of Google Chromecast and Apple TV, which already come with attached storage, but need to be interconnected with neighbours). In this kind of architecture, content can be found using simple routing algorithms (like flooding, CESU [3]).

In this paper, we consider the following problem: *At each cell, which content and where to optimally place it in order to achieve the largest energy savings and smallest delay*. Two foremost concerns to be solved are *what* and *where* to cache. We address these questions at a per-user level: at each user home, we try to cache contents which are of his/her interest. However, there may be neighbouring users also interested in the same content. The option of storing multiple copies at each neighboring user is a more expensive solution (more cache space and more delay) than accessing to the content from the neighbour. To solve this problem, we rely on centrality measures to maximize content shareability.

## 2. RELATED WORK

Throughout the years, researchers have focused on how to mitigate the effect of the media traffic which had an exponential growth over the last decade by caching them in in-network nodes [4]. In this paper, we propose a strategy based on game theoretic centrality that minimizes latency by putting the content close to the interested users(s) so that a content is available to a user within a few hops. Centrality is found to be one of the most promising solutions to this kind of content placement and sharing problem [5], [6]. In [5] the authors explored the advantages of centrality-based caching in Small Base Stations (SBS) scenario and compared performance results among four centrality measures - degree [7], closeness [8], betweenness [9] and eigenvector [10]. In [6] the authors also proposed a caching scheme based on betweenness centrality and showed that their solution can achieve better gain both in synthetic and real topologies. In [11] the authors proposed a new community detection inspired proactive caching scheme for device-to-device (D2D) enabled networks and used Eigenvector Centrality measure to locate the influential users in the community structure. But in none of the works user preference was taken into account. In this paper we use game theoretic centrality [12] to design caching strategy for connected edge networks only. We identify the most suitable node for content placement so that maximum number of users can access the same content with minimum latency.

<sup>1</sup>For example, in the UK, the average home broadband speed in 2017 was 46.2 Mbps <https://www.ofcom.org.uk/research-and-data/telecoms-research/broadband-research/home-broadband-performance-2017>

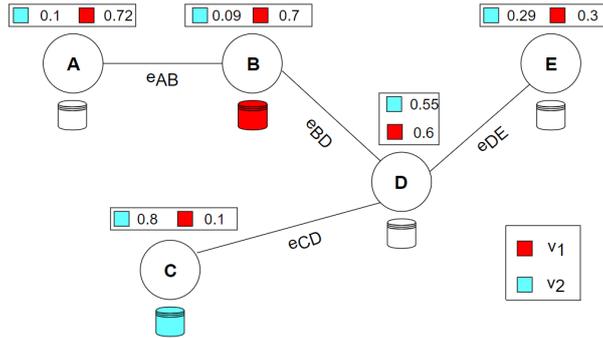


Figure 1: Network example with 5 interconnected users

### 3. PROBLEM STATEMENT AND PROPOSED SOLUTION

Let us define a graph  $V = (N, E)$  where the nodes  $N$  are the user's equipment (e.g., ISPs modem) and the edges  $E$  are Wi-Fi links with some capacity. The capacity of each Wi-Fi link can be calculated based on the distance and existing interference, applying some realistic pathloss model [13]. Given this graph, the content could be placed at the most influential node, based on different centrality measures (e.g., degree centrality, betweenness centrality, eigenvector centrality, closeness centrality). We propose in this work, to use the game theoretic centrality in order to take also into account the past interests of each user.

A *normal form game* (for a more detailed understanding, please refer to [14]) consists of three components -

- A set of players  $N = \{1, 2, \dots, n\}$ .
- A set of pure strategies  $\{S_1, S_2, \dots, S_n\}$ .
- A set of payoff functions  $\{pf_1, pf_2, \dots, pf_n\}$  each assigning a value to each combination of chosen strategies.

In a cooperative game, the players try to come to an agreement, and they also have a choice to bargain with each other, so that they can gain maximum benefit.

In a cooperative game, any number of players can form a coalition  $S$ . A payoff function  $pf(\cdot)$  can be defined to assign an award value  $pf(S)$  to each coalition. The importance of each node in the coalition is measured by calculating the Shapley Value [15]. For each player, this value is calculated in each possible coalition and the sum of all these values tells us how influential the player is. For any player  $i$ ,  $i \in \{1, 2, \dots, n\}$ , the Shapley value can be calculated as

$$\zeta_i(pf) = \sum_{S \subseteq N, i \notin S} \frac{|S|!(|N| - 1 - |S|)!}{|N|!} \cdot [pf(S \cup \{i\}) - pf(S)] \quad (1)$$

The game theoretic centrality is then defined as  $C_{game-theoretic}(i) = \zeta_i(pf)$ . The node with the highest Shapley value is considered as the most influential node in the network.

a) *Input*: Let us apply this metric to the network example depicted in Figure 1. The nodes  $N = \{A, B, C, D, E\}$  denote users that watch content (e.g., videos) from the same streaming service (e.g., Netflix). Let us assume that their location guarantee Wi-Fi connectivity as shown with edges  $E = \{e_{AB}, e_{BD}, e_{CD}, e_{DE}\}$ . A user  $n_i$  can only share content via Wi-Fi with another user  $n_j$  if there exists an edge  $e_{ij}$  between them. Apart from the user connectivity, the problem relies on the concept of *User Interest Probability (UIP)*,  $p_{u,v}$ , which is defined as the probability that a certain user  $u$  will watch a certain content  $v$  within a time period  $(0, T]$ . This value can be calculated based on user's previous watching history ( $t < 0$ ), the video's popularity and many other factors. In this example, let us assume the contents are:  $v_1, v_2$  and the *UIPs* are:  $p_{A,v_1}=0.95$ ,  $p_{B,v_1}=0.85$ ,  $p_{C,v_1}=0.2$ ,  $p_{A,v_2}=0.1$ ,  $p_{B,v_2}=0.09$ , etc. as shown in Figure 1.

b) *Payoff calculation for each possible coalition*: All the nodes can have the strategy of forming a coalition  $S$  by cooperating with other nodes, which results in a subgraph formed by the nodes retaining the same properties from the original graph. A coalition can be formed by any number of nodes in the graph. In our example, the possible coalitions are  $\{\emptyset\}, \{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{A, B\}, \{A, C\}, \dots, N$ . The payoff function  $pf(\cdot)$  assigns an award value  $pf(S)$  to each coalition  $S$  is given by the game policy. The payoff for an absence of coalition  $pf(\emptyset)$  is 0. In our game, the payoff function of a coalition is equal to the sum of all *UIPs* of the users of the coalition when they are able to share content among each other; otherwise, the payoff function is null. Mathematically, it can be expressed for a certain video  $v$  as

$$pf_v(S) = \begin{cases} \sum_{i \in S} p_{i,v} & \text{if } S \text{ is connected} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where  $p_{i,v}$  is calculated at the time of content placement only (i.e., at  $t = 0$ ). The user interest probability values vary with the content, and so do the payoff values. In our example, for all the coalitions, the payoff function

is calculated using the payoff definition of Eqn. 2. For example,  $pf_{v_1}(\{A\}) = 0,72$ ,  $pf_{v_1}(\{A, B\}) = 1,42$ ,  $pf_{v_1}(\{A, C\}) = 0$  (as  $A$  and  $C$  are not a connected subgraph),  $pf_{v_1}(\{A, B, D\}) = 2,02$ , and so on.

c) *Shapley value calculation of each node*: Based on this payoff function, the Shapley value of each player is then calculated using Eq. 1. In our example, the Shapley values for each node and for each content are calculated. For  $v_1$ , the Shapley values are:  $\zeta_A(pf) = 0.281$ ,  $\zeta_B(pf) = 1.081$ ,  $\zeta_C(pf) = -0.081$ ,  $\zeta_D(pf) = 1.059$  and  $\zeta_E(pf) = -0.035$ . As  $B$  has the highest Shapley value, it is the selected location for  $v_1$ . In the same way, for  $v_2$ , the Shapley values are:  $\zeta_A(pf) = -0.178$ ,  $\zeta_B(pf) = 0.471$ ,  $\zeta_C(pf) = 0.991$ ,  $\zeta_D(pf) = 0.344$  and  $\zeta_E(pf) = 0.060$ , and  $C$  is selected to store  $v_2$ .

Let us compare the proposed strategy with the state-of-the-art centrality measures: degree [7], closeness [8], betweenness [9] and eigenvector [10]. These measures do not take into account the user preferences for different videos and hence, for this example, both contents  $v_1$  and  $v_2$  will be cached at  $D$ . But as game theoretic centrality takes into account the user preferences,  $v_1$  and  $v_2$  are cached at  $B$  and  $C$  respectively (as shown in Fig. 1).

## 4. PERFORMANCE EVALUATION

### 4.1 Performance metrics

Let us first define the performance metrics that will be used to evaluate the different placement strategies:

a) *Energy savings*: We use the energy consumption model proposed by Valancius in [16]. Let us denote  $E_1$  as the energy consumed while fetching the item from the server and  $E_2$  as the energy consumed for each sharing (that is, from any user to the node where its content is placed) and  $E_1 > E_2$ . If from  $U \subseteq N$  interested users,  $m$  users fetch the item from server and  $n$  users share from them ( $U = n + m$ ), then the energy savings will be  $1 - \frac{(m \cdot E_1 + n \cdot E_2)}{N \cdot E_1}$ .

b) *Average hopcount per sharing  $\bar{h}$* : This is a measurement of how close the proposed strategy has been able to put the content with respect to interested user(s). The lower  $\bar{h}$  is, the lesser number of hops is required by a node to get the wanted content, i.e., the closer the content is to the interested user. If user  $n_1$  gets the content via  $h_1$  hops,  $n_2$  gets the content via  $h_2$  hops, . . . ,  $n_U$  gets the content via  $h_U$  hops, then  $\bar{h}$  is defined as their average, i.e.,  $\bar{h} = \frac{h_1 + h_2 + \dots + h_U}{U}$ .

c) *Green users*: Let us define as green user as the one who gets the content in at most one hop (i.e., either caches the content or gets it directly from its direct neighbor). A green user does not utilize any link capacity from other users. The larger the number of green users, the better is the utilization of storage space and link capacity in the network.

These three metrics have been measured with respect to the *predictor accuracy*, which indicates how correctly the user preference is guessed. The higher the predictor accuracy is, the more accurately it can predict if a user is going to watch a content and more accurate are the values of the *UIPs*.

### 4.2 Comparison of game theoretic centrality with other centralities metrics

Let us now consider a more generic scenario of  $N = 10$  users, who are connected with Wi-Fi and can share content among them. A user  $n_i$  has a Wi-Fi signal of bandwidth  $b_i$ . Let us assume that all users have the same bandwidth  $b_i = B$ . Let  $b_{ij} < B$  denote the available bandwidth between the users  $i$  and  $j$  for sharing content. The minimum bitrate required to play a video is denoted by  $k$ . Both  $b_{ij}$  and  $k$  are normalized with respect to  $b$ .

For our simulations, we generate a sparsely connected graph with  $N = 10$ . We randomly assign a value between 0 and 1 to each of  $b_{ij} \forall i, j \in N$  which signifies the link capacity. This value is set before the start of the simulation and does not change throughout the simulation. Now, for each content  $c$ , we assign another randomly chosen value between 0 and 1 that signifies the bitrate  $k$  of  $c$ . For all  $i, j \in N$ , if the value of  $b_{ij}$  is greater than  $k$ , then we assume that  $c$  can be shared between them and an edge is created between  $i$  and  $j$ . Also, for each user  $i$ , we randomly assign user interest probability values  $p_{i,c} \forall i \in N$  and do a weighted random selection to decide which user(s) will watch  $c$ .

We run the simulation for a confidence interval of 90%, each time randomly assigning values to  $k$  and  $p_{i,c}$ . We measure the energy savings, average hopcount per sharing and green users with respect to predictor accuracy. As this is not a real life dataset, we do not know precisely the size of the content universe and thus we do not perform measurements with respect to storage. After we run all the experiments, then the average is taken and error bar is also plotted. We consider the existing centrality measures to find the most influential user in the cluster where the content can be cached and compare the performance with our proposed strategy.

**Energy savings** We measure the energy savings while varying the predictor accuracy. From Fig. 2a we see that following our proposed strategy offers more than 80% savings in energy can be obtained, which is more than the energy savings obtained by any of the state-of-the-art centrality measures.

**Average hopcount per sharing** The lower is the value of the average hopcount per sharing, the closer the content has been put to the interested user(s). From Fig. 2b we see that for our proposed strategy, this value is much smaller than the other ones. Following our strategy, a user is able to find content within 2-3 hops.

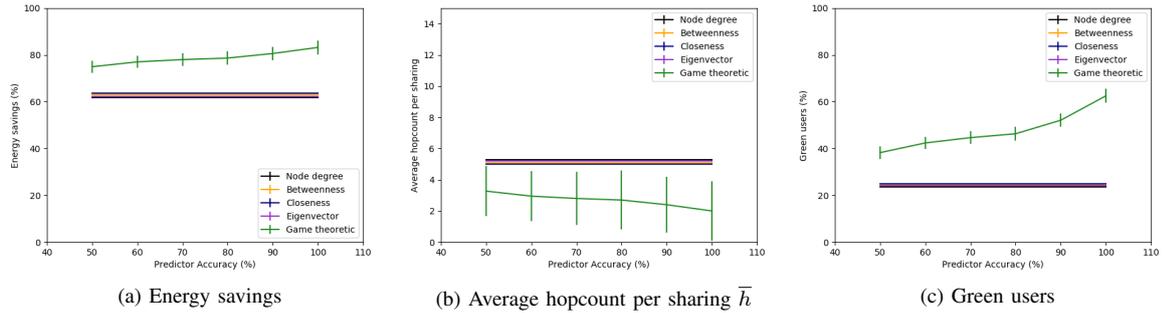


Figure 2: Performance for varying parameters

**Green users** The more is the number of green users, the better is the resource utilization in the network. From Fig. 2c we see that for other centrality measures the number of green users is around 25%, while using our strategy, the number of green users lies between 40-70%.

## 5. CONCLUSION

In this paper we proposed a proactive caching strategy for a group of users who are able to share content among themselves using Wi-Fi. We used game theoretic centrality to find the most suitable node in the network for caching a content and discussed the advantages of using proposed approach over the existing ones. Despite having the dependence on the accuracy of the predictor regarding calculation of the  $UIP$  values, on an average a user could find the content he was looking for within 2-4 hops in a sparsely connected graph. We define the payoff function so that content is placed close to the interested user(s) and sharing becomes maximum. To optimize different parameters like storage or link utilization, payoff function can be defined differently and the performance can be studied.

## REFERENCES

- [1] F. Boccardi *et al.*, "Five disruptive technology directions for 5G," *IEEE COMMAG*, vol. 52, no. 2, pp. 74–80, 2014.
- [2] A. Raman *et al.*, "Wi-Stitch: Content delivery in converged edge networks," in *Proc. of the Workshop on Mobile Edge Communications*, pp. 13–18, ACM, 2017.
- [3] A. Banerjee, B. Banerjee, A. Seetharam, and C. Tellambura, "Content search and routing under custodian unavailability in information-centric networks," *Computer Networks*, vol. 141, pp. 92–101, 2018.
- [4] G. Zhang *et al.*, "Caching in information centric networking: A survey," *Computer Networks*, vol. 57, no. 16, pp. 3128–3141, 2013.
- [5] E. Baştuğ *et al.*, "Centrality-based caching for mobile wireless networks," in *1st KuVS Workshop on Anticipatory Networks*, 2014.
- [6] W. K. Chai *et al.*, "Cache less for more in information-centric networks," in *Int. Conf. on Research in Networking*, pp. 27–40, Springer, 2012.
- [7] L. C. Proctor C.H., "Analysis of sociometric data. In: Holland," *Research Methods in Social Relations*, pp. 561–586, 1951.
- [8] A. Bavelas, "Communication patterns in task-oriented groups," *The Journal of the Acoustical Society of America*, vol. 22, no. 6, pp. 725–730, 1950.
- [9] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, pp. 35–41, 1977.
- [10] P. Bonacich, "Power and centrality: A family of measures," *Am. journal of sociology*, vol. 92, pp. 1170–1182, 1987.
- [11] A. Said *et al.*, "Proactive Caching at the Edge Leveraging Influential User Detection in Cellular D2D Networks," *Future Internet*, vol. 10, no. 10, p. 93, 2018.
- [12] R. Lindelauf *et al.*, "Cooperative game theoretic centrality analysis of terrorist networks: The cases of jemaah islamiyah and al qaeda," *European Journal of Operational Research*, vol. 229, no. 1, pp. 230–238, 2013.
- [13] P. Series, "Propagation data and prediction methods for the planning of short-range outdoor radiocommunication systems and radio local area networks in the frequency range 300 MHz to 100 GHz," tech. rep., ITU, Tech. Rep. ITU-R, 2015.
- [14] S. Tadelis, *Game theory: an introduction*. Princeton University Press, 2013.
- [15] L. Shapley, "A value for n-person games," *Annals of Mathematics Studies*, vol. 28, pp. 307–317, 1953.
- [16] V. Valancius, N. Laoutaris, L. Massoulié, C. Diot, and P. Rodriguez, "Greening the internet with nano data centers," in *Proceedings of the 5th international conference on Emerging networking experiments and technologies*, pp. 37–48, ACM, 2009.