

Social Networking and Caching Aided Collaborative Computing for the Internet of Things

Yutong Ai, Li Wang, *Senior Member, IEEE*, Zhu Han, *Fellow, IEEE*,
Ping Zhang, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

Abstract—The Internet of Things (IoT), as the latest information technology, is expected to support the ubiquitous connectivity of anything, anytime and anywhere. Since numerous IoT devices will be connected to the cellular networks, the conflicting demands of ultra-low latency, massive connectivity and high reliability impose significant burdens on the operational wireless networks. Exploiting various temporarily unused computational and storage resources by invoking cooperative caching and computing has the potential of supporting these demanding applications. Therefore, the burdens imposed on wireless networks by the IoT applications can be mitigated by the proposed social networking and caching assisted collaborative computing IoT framework.

Index Terms—Internet of Things (IoT), Social Networking, Caching, Computing, and Non-Orthogonal Multiple Access

I. INTRODUCTION

The Internet of Things (IoT) is a revolutionary communication paradigm conceived for connecting a plethora of devices together [1], relying on communication, caching and computing in wireless networks. The family of IoT application scenarios can be generally divided into two broad categories. In *massive IoT* applications, sensors typically report to the central processor on a regular basis, whilst relying on low-cost devices with low energy consumption. Their application examples include smart buildings, logistics, tracking and fleet management. In *critical IoT* applications there are high demands in terms of reliability, availability and low latency. Critical IoT applications include remote health monitoring, traffic safety and control, industrial applications and control, remote manufacturing, training, surgery, etc [2]. It is predicted that the IoT will include over 25 billion units by 2020 [3]. Naturally, this dramatic proliferation of the emerging IoT technologies will impose significant challenges on the design of wireless networks.

Y. Ai, L. Wang (*corresponding author*) are with the School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, China. Email: {ytai, liwang}@bupt.edu.cn.

Z. Han is with Department of Electrical and Computer Engineering, University of Houston, TX, USA. Email: zhan2@uh.edu.

P. Zhang is with School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. Email: pzhang@bupt.edu.cn.

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K. Email: lh@ecs.soton.ac.uk.

L. Hanzo would like to acknowledge the financial support of the EPSRC projects EP/N004558/1, EP/P034284/1, of the Royal Society's GRFC Grant as well as of the European Research Council's Advanced Fellow Grant QuantCom.

In IoT networks, there are diverse communication equipment having various computational and storage capability. In order to further improve the efficiency of communication in this treatise we propose to invoke cooperative caching and computing. The devices in the wireless network may engage in activities beyond their own capability. By cooperating with nearby devices, Ding *et al.* [4] designed a cognitive network for efficiently exploiting the temporarily unused spectral resources. By contrast, our goal is to exploit the nearby devices for cooperative caching and computing by relying on social networking.

Traditional wireless networking mainly considers the wave propagation characteristics and the physical locations of users as well as their busy/idle status, but not their social characteristics. To fill this gap, our proposed framework exploits the social relationships of users such as the typical frequency and duration of their contact, because cooperating with nodes having a high-quality link improves the system performance attained [5]. By jointly exploiting the physical and social characteristics, a more powerful system may be designed [6].

In order to further enhance the performance of our framework, the beneficial selection of multiple access techniques constitutes another vital factor, which should be taken into account. Multiple access techniques may be broadly classified into orthogonal and non-orthogonal approaches. In Orthogonal Multiple Access (OMA), the signals of different users do not directly interfere with each other. By contrast, Non-Orthogonal Multiple Access (NOMA) allows overlapping of the user-signals in the time-, frequency-, spatial-, power- or the code-domain, with the motivation of enhancing the throughput of OMA [7] [8]. In most of the existing wireless networks, the radio resources, including the time and frequency, are orthogonally allocated to different devices for their data transmission. However, in the IoT a large number of devices must be supported, hence using OMA techniques is less practical. Therefore NOMA has been identified as a promising technique of supporting numerous IoT devices.

Additionally, a compelling combination of social networking, cooperative caching and computing may be invoked for supporting IoT applications. Hence in Section II we commence by outlining our social networking and caching aided collaborative computing system framework proposed for IoT networks. Given this cross-layer architecture, in Section III we analyze the social considerations in the proposed framework. Next, we introduce coded caching and illustrate its advantages as well as challenges in Section IV, whilst our NOMA solution advocated is analyzed in Section V. Furthermore, the above-

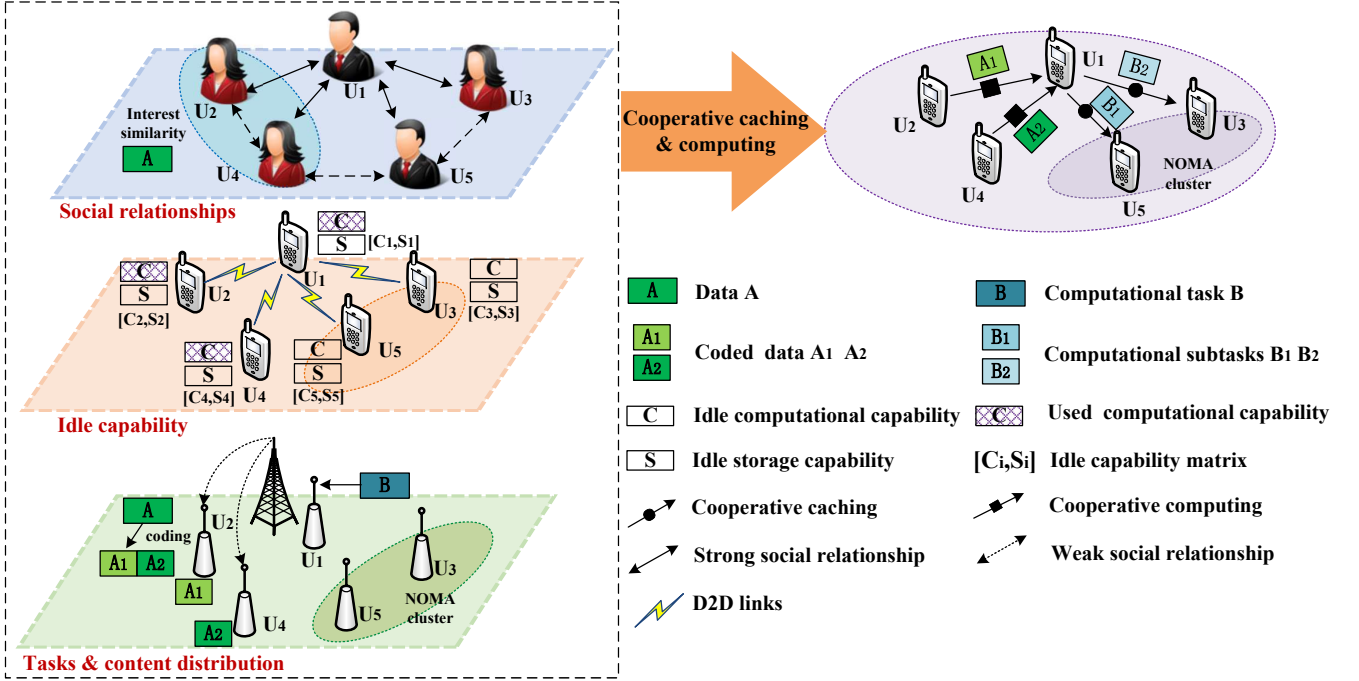


Fig. 1. An illustration of the proposed system framework.

mentioned techniques are studied in Section VI. Finally, we conclude in Section VII.

II. THE PROPOSED SOCIAL NETWORKING AND CACHING AIDED COLLABORATIVE COMPUTING FRAMEWORK

Sophisticated IoT applications require large storage and substantial computational capability at a low cost. To resolve these design-conflicts, we propose a social networking aided cooperative framework to exploit the temporarily unused computational and storage capability of idle social contacts both for caching and computing.

An illustration of the proposed framework is shown in Fig. 1 and summarized as follows.

- 1) **Information Collection:** The social relationships of nodes and their cooperative inclination is surveyed, with special emphasis on their willingness to share their idle storage and computational capability. The information collected is delivered to the base station (BS) for exploiting it in its coverage area. The corresponding information shown in Fig. 1 is as follows:

- 1) the social relationships between U_1 and U_2 - U_5 are strong, hence U_2 - U_5 are willing to cooperate with U_1 ;
- 2) U_2 and U_4 have similar interests in data A and also have some idle storage capability;
- 3) the tasks of U_1 are beyond its computational capability;
- 4) U_3 and U_5 have idle computational capability, which can be exploited;

Hence the 'idle capability' matrix is introduced to quantify the users' capability of supporting caching aided collaborative computing.

- 2) **Data Content & Computational Task Distribution:** During content caching, data A is cached both into U_2 and U_4 , since they have idle storage capability and share a common interest in data A . When U_1 requires data A for executing the computational task B , it fails owing to the exhaustion of its own computational capabilities. Therefore, U_1 has to gain access to data A first and then resort to the social contacts' assistance for carrying out the computational tasks required.
- 3) **Caching Aided Collaborative Computing Policy:** By invoking cooperative caching, data A can be acquired from U_2 and U_4 . Having obtained data A , it turns out that U_1 still cannot process data A due to its limited computing power. Hence U_1 partitions the related computational task B into two subtasks and allocates them to U_3 and U_5 for cooperative computing. After U_3 and U_5 complete their tasks, the computational results are returned to U_1 . To improve the caching efficiency, data A is split into A_1 and A_2 first and then transmitted. In this context, U_3 and U_4 may form a NOMA cluster for their communication. Therefore, the final caching aided collaborative computing scheme is shown at the right of Fig. 1.

Naturally, the consideration of the associated social relationships also significantly affects the efficiency of cooperation. Therefore, a cross-layer operation aided cooperative caching and computing scheme should be designed. In the following section, we will analyze the impact of both the social relationships as well as of the advantages of coded caching and of the NOMA framework advocated.

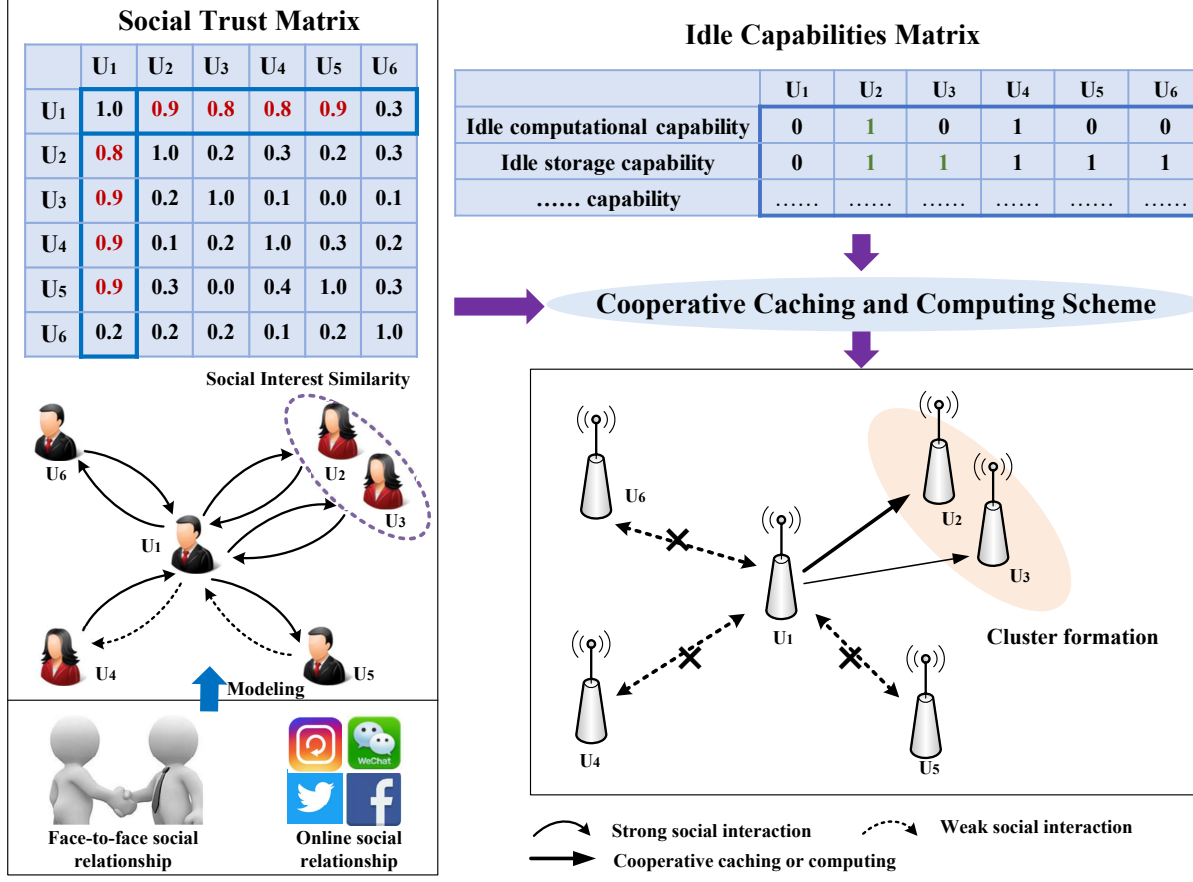


Fig. 2. An illustration of the proposed framework based on social relationships.

III. SOCIAL BEHAVIOR INFERENCE AND IMPACT

In IoT networks, the activity of the devices is intrinsically linked to the human behavior, after all, the IoT applications are developed to satisfy the human requirements. The social relationship furnishes information for the IoT devices. The social relationships may manifest themselves in terms of face-to-face social relationships and online social relationships [9]. Face-to-face social relationships are mainly established by localized interactions, while the online social relationships rely on Facebook, Instagram, LinkedIn and Twitter, etc. In these two types of social relationships, the following three aspects will be considered for reliable communication in our framework.

A. Social Modeling Based on Social Behaviors

- 1) **Social Interaction:** direct communication can indeed be realized, but high-integrity communication cannot be guaranteed because of the mobility of nodes, where the social interaction can be characterized by the contact duration and contact frequency between users [5]. Specifically, the data is deemed to be successfully transmitted to others, if the transmission process is accomplished within the contact duration.

- 2) **Social Trust:** social trust plays a vital role in determining the level of altruism amongst users, quantified by a non-negative number $s \in [0, 1]$, where $s = 1$ indicates unconditional availability and $s = 0$ indicates no availability. The social trust between each pair of users can be quantified by the matrix shown in Fig. 2.
- 3) **Social Interest Similarity:** users in a certain area or during a certain period are likely to have interests in the same content. Social interest similarity is hence a metric quantifying how similar the interests of two users are, which may be used for partitioning the network into user-groups for improving the performance by content sharing. Instead of repeatedly requesting the content of interest from the BS, the user may request it from the users in the same interest group.

In our framework, the social relationships are considered to be the driving factors of cooperation. On the one hand, nodes having certain needs tend to seek the nearby nodes having both available resources and common interests, who are willing to help them. On the other hand, the requested nodes will decide whether to offer help based on the reputation of the requesting nodes.

B. Social Influence Case Study

Fig. 2 shows the influence of the social relationships on our proposed scheme. Here, we assume that U_1 has exhausted its capability and hence needs cooperation for completing its tasks.

First, the associated social relationships are surveyed based on the users' reports. When considering the social trust matrix from the perspective of U_1 , the social trust indices of U_2 - U_5 are 0.9, 0.8, 0.8, and 0.9, respectively, meeting the requirement of U_1 in terms of social trust. By contrast, U_6 is not sufficiently trustworthy. Additionally, U_2 - U_5 all believe that U_1 is dependable and hence they are willing to aid U_1 . However, U_4 and U_5 still cannot assist U_1 , since their social relationships with U_1 are relatively weak. Bear in mind that U_1 's trust in U_2 is not necessarily reciprocated at the same level, but the social interaction has to be also bidirectional.

The associated capability matrix is also shown in Fig. 2. For the sake of simplicity, we use "1" to indicate that the user does have idle capability, otherwise we write "0". It is clear that U_2 and U_3 do have idle capability. According to the above-mentioned information, a simple cooperative caching and computing scheme can be conceived, where U_1 relies on cooperative caching or computing provided by U_2 , but only on cooperative caching provided by U_3 . Based on their similar interest, U_2 and U_3 may form a cooperative cluster to perform cooperative caching. It can be readily seen that the social relationships are important for beneficial candidate selection, when invoking caching aided collaborative computing.

Furthermore, when the data content or the computational tasks have to be partitioned into several parts, the specific partitioning has to be carefully considered. Apart from the capability of users, their social relationship is another non-negligible aspect. Social trust can be exploited for beneficially designing the content/computational task allocation schemes. For example, more popular data content can be cached in more dependable nodes and similarly, more computational tasks may be allocated to more reliable nodes, such as U_2 in Fig. 2.

Social relationships constitute an important consideration in our proposed framework, but it may be a challenge to establish an accurate mathematical model of the social relationships amongst users. Machine learning can be used for surveying the complex social relationships and for characterizing the social behaviors [10].

IV. EXPLOITING CODED CACHING IN THE PROPOSED FRAMEWORK

By prefetching popular content during off-peak times for the devices, wireless caching is capable of alleviating the peak-hour network congestion, especially for delay-tolerant applications. Again, IoT devices may request desired content from each other [11] [12] with the aid of cooperative caching amongst them. Therefore, caching-aided networks have the potential of reducing the backhaul load, in turn reducing the end-to-end access delay and increasing the peak rate [13]. In multi-user caching scenarios the social relationships developed in the previous section can be exploited to form both cooperative clusters and the NOMA-pairs to be developed in the next section.

Furthermore, coded caching exploits that each content of interest can be partitioned into multiple content fragments, which may be cached in different devices, hence further increasing the caching efficiency [14]. When a user sends a content request, nearby devices which have cached the different content fragments of the content items requested are capable of delivering the desired content fragments. For example, the maximum distance separable (MDS)-coded caching schemes [11] split each content item into n parts and encode them into $N(N > n)$ coded pieces. Once the original n coded fragments have been collected, the original content can be recovered, which beneficially reduces both the content download time and the backhaul rate. Apart from the above-mentioned advantages, coded caching can also be exploited for caching aided collaborative computing.

Cache-enabled wireless networks operate in two phases in general, namely the content placement and content delivery. During the content placement, the collection of information concerning the user-capability and social relationships will be performed first. As shown in Fig. 3, based on the information collected, the BS distributes the content to the most suitable zones, where the users have common interests. According to the social relationships of nodes, popular content will be stored by the trusted users to increase the successful content sharing probability. With the aid of coded caching the content delivery, cooperative caching and collaborative computing become more flexible. In Fig. 4 *Case 1*, U_3 can download content A from U_1 and U_2 . Therefore, content A can be readily retrieved from the close-by users instead of the BS to avoid excessive loading for the BS. With the advent of coded caching, content A is not necessary fetched from a single user. This is practically beneficial, when content A is excessive or when a cooperator is only prepared to offer limited assistance.

To apply coded caching in our proposed framework, a range of practical issues have to be solved. Because of the limited computational capability of IoT devices, the coded strategy associated with a low complexity is indeed compelling. Furthermore, both the user's interest in the content and their social relationships are time-variant. Hence the cooperative caching scheme has to be adapted accordingly.

V. NOMA AIDED SOCIAL NETWORKING FOR CONTENT DELIVERY AND COMPUTATION OFFLOADING

In the fully-fledged IoT networking of the near future numerous devices must be supported within a small region. Hence it is unrealistic to have an exclusive orthogonal channel for each device. Fortunately, NOMA allows multiple devices to use the same resource-slot, hence improving the spectral efficiency of conventional OMA.

A. Basic Concept of Power-Domain NOMA

For simplicity, we assume that there are two users in a NOMA cluster, each having a single transmit and receive antenna. The basic concept of NOMA is shown in Fig. 4. According to [15], the NOMA cluster may support a user having higher transmit power and a poorer channel condition as well as a lower-power user having better channel conditions.

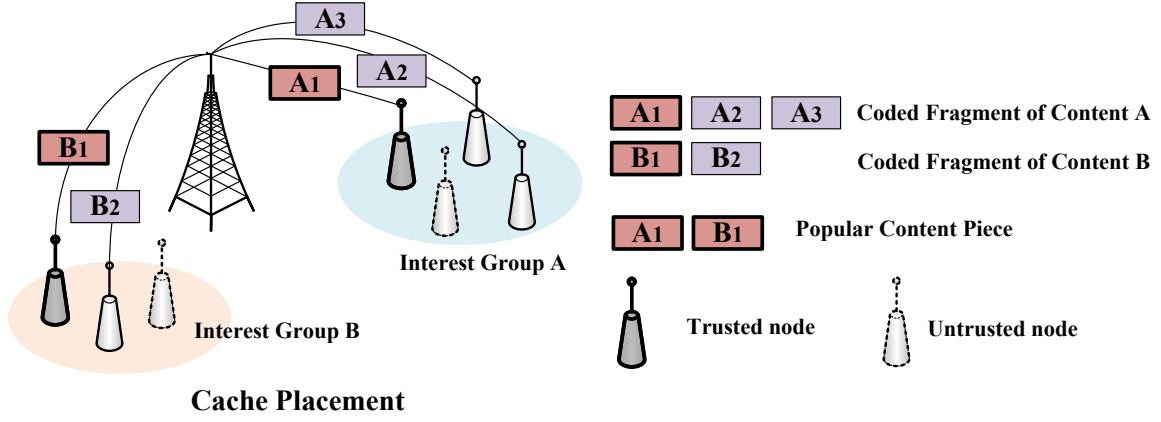


Fig. 3. An illustration of cache placement based on social relationships.

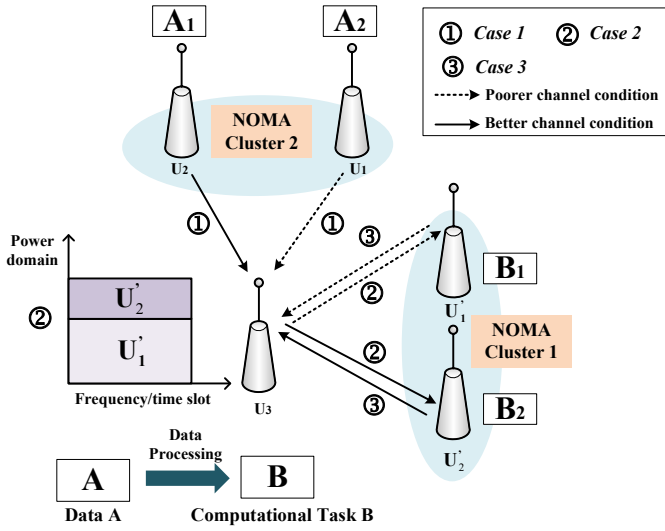


Fig. 4. Exploiting coded caching and NOMA in our proposed framework.

Let us assume that U_1 is assigned a higher and U_2 a lower transmit power. During transmission in the downlink, the signals of U_1 and U_2 are superposed first and they are transmitted over the same time- and frequency-slot. U_1 directly decodes its own higher-power signal which is only slightly contaminated by the lower-power U_2 . At U_2 the higher-power U_1 signal is not required, but since the higher-power signal can be more reliably detected, it is this one, which is detected first and then remodulated, followed by its subtraction from the composite $(U_1 + U_2)$ signal to generate the noise-contaminated U_2 signal.

B. NOMA for Heterogeneous IoT

As mentioned, a substantial benefit of NOMA invoked for the IoT is that the devices transmit in the same time- and frequency-slot, which leads to efficient exploitation of the limited spectral resources. Additionally, when the channel gain between U_1 and the BS is significantly different from that of the BS- U_2 link, the performance gain of NOMA over OMA can be significant. Since the battery of IoT devices is limited,

the energy consumption is also a crucial issue to be considered. NOMA is capable of enhancing the energy efficiency.

Specifically, in Fig. 4 *Case 1*, U_1 and U_2 form a NOMA cluster. The content A is partitioned into two fragments, which are transmitted simultaneously to U_3 through the same channel. Similarly, U_3 can also delegate computational subtasks to U_1' , U_2' and then fetch the computational results through the same channel, as seen in Fig. 4 *Case 2* and *Case 3*.

Although the NOMA concept brings about several advantages in our proposed framework, there are still several challenges as for its practical application in the proposed IoT framework. Firstly, our proposed framework considers both physical and social system characteristics, hence the NOMA cluster formation and resource allocation schemes are more complex. But IoT devices cannot afford complex calculations, hence low-complexity yet efficient allocation algorithms are required. Secondly, when the number of cooperative users in a NOMA cluster is increased, the spectral efficiency can be enhanced. However, having more cooperative NOMA users mapped to the same time- and frequency-slots will cause more interference and makes the process of successive interference cancellation (SIC) more challenging for both the cooperative users and for the BSs.

VI. EXAMPLE SCENARIO AND ITS PERFORMANCE

In this section, we characterize the performance of the proposed framework. To illustrate the effects of social relationships, we introduce a non-negative number $S_{i,j}$ to model the social relationships between users i and j where having a higher $S_{i,j}$ means a stronger social relationship. When $S_{i,j}$ is sufficiently high, caching aided collaborative computing can be relied upon. The simplest possible NOMA cluster supporting as few as two users is exploited for transmitting the desired content or computational task. We assume that every content file is partitioned into two independent fragments and that two users of the NOMA cluster will cache a single fragment of content. Similarly, a specific computational task may be divided into two subtasks for cooperative computing by the NOMA cluster. The cooperation is deemed successful, when both cooperative users complete their caching aided collaborative computing task. During the computing process,

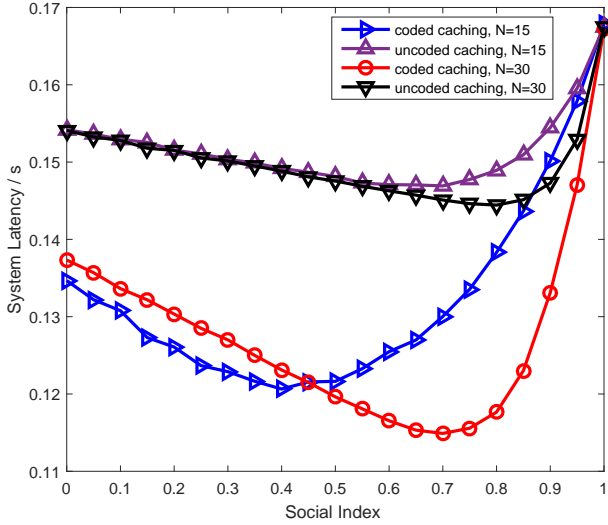


Fig. 5. System latency of the proposed framework relying on coded caching versus uncoded caching in a multi-user scenario, where the total number of users is $N = 15$ and 30 .

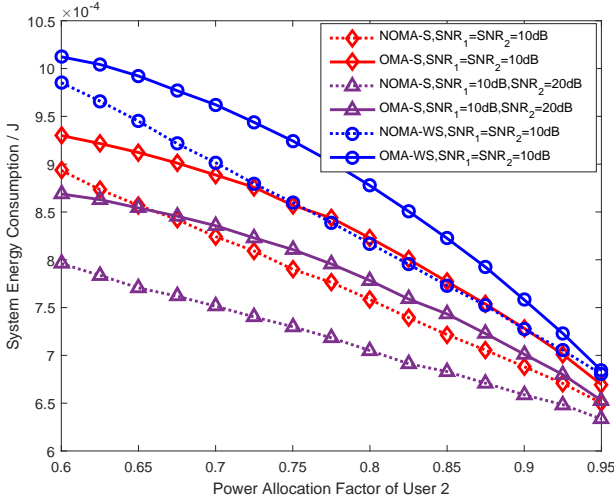


Fig. 6. Energy consumption of the proposed framework relying on NOMA exploited and OMA in a two-user scenario.

we assume that the users have different CPU clock frequencies of $\{1.0, 1.2, 1.4, 1.6\}$ GHz. The complexity of the computational tasks is quantified in terms of the number of bits to be processed given by $\{1.0 \times 10^5, 1.5 \times 10^5, 2.0 \times 10^5\}$ bits, where the number of CPU cycles required per bit obeys the uniform distribution in the range of $[500, 1000]$ cycles/bit. The total bandwidth is 10MHz, which is divided into 10 subbands for data transmission. In OMA, every subband supports only one of the cooperating users, while NOMA supports two cooperating users in each subband.

Fig. 5 shows the system's latency quantified in terms of the total time required by the caching aided collaborative computing task. The social index represents the requirements concerning the users social relationships. A higher social index indicates a user's preference for a stronger social tie, higher social trust and more similar interest. Observe in Fig. 5 that the social relationship is an important factor in terms of prede-

termining the performance improvements attained, especially for the situation where more users are involved. However, relying on an excessively high social index makes it hard to perform cooperative caching and computing, hence resulting in a degraded system performance. The system performance attained may be further enhanced, when coded caching is invoked, which increases the speed of data sharing.

Fig. 6 shows the energy consumption of a two-user scenario dealing with different computational tasks or contents. It is clearly seen that NOMA imposes a lower energy consumption, since NOMA enhances the transmission efficiency. Furthermore, when the difference of channel conditions becomes more pronounced, significant improvements can be achieved by NOMA over OMA.

VII. CONCLUSIONS

A social networking and caching aided collaborative computing framework was proposed for IoT networks by considering the interplay between the physical system and social characteristics. We have quantified the impact of social relationships on our framework. To further enhance the performance, both coded caching and NOMA techniques have been incorporated. Finally, our numerical results have characterized the efficiency of the proposed framework.

REFERENCES

- [1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22-32, Feb. 2014.
- [2] M. Shirvanimoghaddam, M. Dohler, and S. J. Johnson, "Massive non-orthogonal multiple access for cellular IoT: potentials and limitations," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 55-61, Sep. 2017.
- [3] Ericsson, "Cellular Networks for Massive IoT," tech. rep. Uen 284 23-3278, Jan. 2016; https://www.ericsson.com/res/docs/whitepapers/wp_iot.pdf.
- [4] H. Ding, Y. Fang, X. Huang, M. Pan, P. Li, and S. Glisic, "Cognitive capacity harvesting networks: architectural evolution toward future cognitive radio networks," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 3, pp. 1902-1923, third quarter 2017.
- [5] L. Wang, H. Wu, and G. L. Stuber, "Cooperative jamming-aided secrecy enhancement in P2P communications with social interaction constraints," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1144-1158, Feb. 2017.
- [6] J. Hu, L. L. Yang, H. V. Poor, and L. Hanzo, "Bridging the social and wireless networking divide: information dissemination in integrated cellular and opportunistic networks," *IEEE Access*, vol. 3, pp. 1809-1848, Sep. 2015.
- [7] Y. Liu, Z. Qin, M. ElKashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Non-orthogonal multiple access for 5G and beyond", *Proceedings of the IEEE*, Dec. 2017, to appear.
- [8] Y. Chen, L. Wang, Y. Ai, B. Jiao, and L. Hanzo, "Performance analysis of NOMA-SM in vehicle-to-vehicle massive MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. PP, no. 99, pp. 1-1, 2017.
- [9] H. Zhao, H. Zhou, C. Yuan, Y. Huang, and J. Chen, "Social discovery: exploring the correlation among three-dimensional social relationships," *IEEE Transactions on Computational Social Systems*, vol. 2, no. 3, pp. 77-87, Sep. 2015.
- [10] Z. Ge, Z. Song, S. X. Ding, and B. Huang, "Data mining and analytics in the process industry: the role of machine learning," *IEEE Access*, vol. 5, pp. 20590-20616, Sep. 2017.
- [11] L. Wang, H. Wu, Y. Ding, W. Chen, and H. V. Poor, "Hypergraph-based wireless distributed storage optimization for cellular D2D underlays," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 10, pp. 2650-2666, Oct. 2016.

- [12] B. Bai, L. Wang, Z. Han, W. Chen, and T. Svensson, "Caching based socially-aware D2D communications in wireless content delivery networks: a hypergraph framework," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 74-81, Aug. 2016.
- [13] G. Paschos, E. Bastug, I. Land, G. Caire, and M. Debbah, "Wireless caching: technical misconceptions and business barriers," *IEEE Communications Magazine*, vol. 54, no. 8, pp. 16-22, Aug. 2016.
- [14] M. A. Maddah-Ali and U. Niesen, "Coding for caching: fundamental limits and practical challenges," *IEEE Communications Magazine*, vol. 54, no. 8, pp. 23-29, Aug. 2016.
- [15] Z. Ding, Z. Zhao, M. Peng, and H. V. Poor, "On the spectral efficiency and security enhancements of NOMA assisted multicast-unicast streaming," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 3151-3163, Jul. 2017.