Advanced Science Letters
Vol. 21, 3548–3551, 2015

Content Caching in ICN using Bee-Colony Optimization Algorithm

Ibrahim Abdullahi[1,*], Suki Arif[1], Suhaidi Hassan[1]

[1]InterNetworks Research Laboratory, School of Computing. Universiti Utara Malaysia, UUM 06010 Kedah.
[*]Department of Mathematics/Computer Science, Ibrahim Badamasi Babangida University, Lapai. Niger State. Nigeria

Information dissemination has recently been overtaken by the huge media-driven data shared across different platforms. Future Internet shall greatly be concerned about pervasion and ubiquity of data on all devices. Information-Centric Network seems the challenging paradigm that aims at guaranteeing the flexibility needed when the data explosion occurs. Caching is thus an option that provides the flexibility that manages data exchange practices. Different caching issues has raised concern about the content flooded all over the Internet. In line with the challenges, Bee-Colony Optimization Algorithm (B-COA) has been proposed in this paper to avail content on the Internet with less referral cost and heavy monopoly of data on hosts. It is believed that the advantages of the grouping and waggle phase could be used to place the contents faster in ICN.

Keywords: Information centric network, cache, content placement, Bee-colony

1. INTRODUCTION

The Information dissemination with the sole purpose of deploying the Internet has been managed in different folds. Among the most benefitting of the platforms is the Internet. Internet has been fully inculcated into the daily affairs of users, educationally, politically, academically, socio-economically and otherwise. However, with the current trend and practices of vital information sharing, the need to maintain data security, integrity, identification, transporting is of the essence.

Internet of host-centrism has benefited the intended design and practices of the years past. User demanding nature of having data and information on every platform nowadays has necessitated the need to project the future practices. This has given birth to the concept of Internet of Things (IoT), Internet of Everything (IoE), Ubiquitous computing, Fly computing, Fog computing and Information Centric Network for the management of data[1,2]. The Information Centric Network also referred to the Content Centric Networking is the Internet paradigm that promises to ameliorate the forecasted problems that might be faced by the current host-centric Internet. The traditional network of Internet Protocol (IP) addressing is faced with the operation of the user specifying the address and location of where data and information is to be fetched. On every iteration that requests are placed, the IP networks need the negotiation of the Domain Name Server (DNS) alongside the Dynamic Host Communication Protocol (DHCP) to guarantee the needed data.

[*]Email Address: ibrahim@internetworks.my

However, ICN on the other hand has been proposed to have the functionality of providing the needed data without specifying the location, and identity of the data or the interest provider in[1,3,4] as shown in Figure 1. The ICN paradigm therefore provides a content-aware technique through its special features. In ICN, data usually referred to as Interests when coming from the user (subscriber) are sent to flood the network. Once an Interest is pushed into the network as subscribed data, the mid-channel routers specially records the Interest and forwards the Interest through its neighbors.

This procedure guarantees through the embedded advantage of not specifying the granter of the information[4]. The ubiquity of data for the future Internet through ICN can therefore be feasible through adequate cache management. The term Interest and Interest packets shall be used throughout this study as the same. However Interests and Interest packets will be used interchangeably in the subsequent sections.

Caching in ICN like any other computer study, is the predictive saving of data/information for temporal use of referral. Caching has been itemized as the leading deliverable to actualize the ICN goal of having information everywhere. Since ICN techniques care less about the address of the publisher when data is hosted. Some ICN approaches have proposed the ways of achieving the content-centrism of information sharing[5]. Among the popular ICN architectures and proposals are NDN, NetInf, PSIRP, DONA etc.

2. RELATED WORKS IN ICN ARCHITECTURE CACHING

NDN[1] also referred to as Content-Centric Network is one of the most actively researched approach of ICN. NDN building block consists of three special data structures, namely: Content Store (CS), Forwarding Information Base (FIB)and the Pending Interest Table (PIT). For the purpose of this study, CS shall be the focal point to exhibit caching. NDN architectural development functions through the initiation of Interest Packets from a subscriber. Request refers to as Interests are initiated by a user from a point close or far from the publisher. Users in NDN do not specify the address of the needed data as against the practice in an IP network. Once the Interests cross the routers, they are automatically recorded in PIT as indexes. Caching practice in NDN is thus seen as an opportunistic practice[1]. NDN caching is therefore referred to as the Leave Copy Down (LCD). It is seen as the caching form that avails information only to the routers crossed during the data search.

NetInf[6,7] is an ICN architecture that benefits from the special dedicated system part called Name Resolution System (NRS). The NRS is a special design aid that records the flooded named data objects (NDO) as

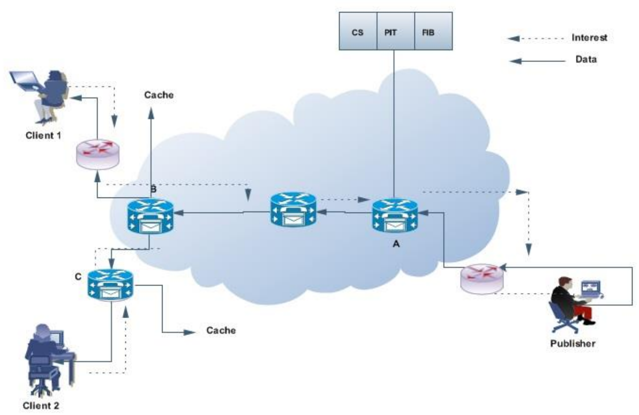

Fig. 1 ICN Overview

publisher registers their available data. Interest packets communicate through the NRS to obtain data lodged by publishers. Caching practice here is time saved as the FIB in routes forwards and locates the data from the publishers. A strong bond between NRS, subscriber and publisher is done through the source locator. In the searches on the network part, Multilevel Distributed Hash Tables (MDHT) are used. The predictor case in NRS is forecasted to be challenged and probable to a single point of failures in its technique.

PSIRP[8] on the other hand is a project also referred to as the PURSUIT. In PSIRP, the subscriber sents out the request and Interest packets in similar fashion as previously itemized in NDN and NetInf. As discussed in NetInf, PSIRP caching and Interest are guaranteed through a special incorporated node known as Rendezvous. The Rendezvous in PURSUIT project is a server-like points that keeps records as caches of the indexes in the network. Rendezvous bridges the subscriber request referred to as registers. The Rendezvous form a mini-network of rendezvous to speed the delivery and fast binding of the registers known as Rendezvous network (RENE)[3]. The challenge in PSIRP can therefore be seen in the linear content delivery.

DONA[9] caching in the content centric network is aided by the resource handlers (RH). The RH(s) perform similar functions as described about PSIRP. The concept of DONA guarantees and sets a Time-To Leave (TTL) for each cached data or Interest. It is thus becoming challenging to researchers in the form of caching that could best benefit the management of the caches. The TTL setting needs to be intelligently set to avoid the content eviction and deletion on needed data in the caches. A probable solution to the aforementioned challenges is to optimize the contents in neighbor routers. Bee-colony optimization algorithm is therefore a viable resource to achieving better cache hit ratios.

3. CACHING IN ICN

A good caching technique is the goal of all ICN deployed projects. Adequate and sufficient cache hits are therefore the necessity for accepting a caching approach[10,11]. In ICN, caching has been categorized into two approaches; namely: On-path caching and Off-path caching.

3.1 Off-path caching

A cache practice in ICN that enables the content evenly distributed to incur more hit ratio. However, as submitted by[2,12], there is the need to provide additional mechanism, routing information and added forwarding strategies. In achieving the off-path, a trade-off is gained on either predicting the benefits of having the options that mitigate evicting popular contents in caches or the high hop crosses. When additional mechanisms are added, hop-counts becomes high. This affects the state of the network and lesser the throughput by negatively altering the latency of the network.

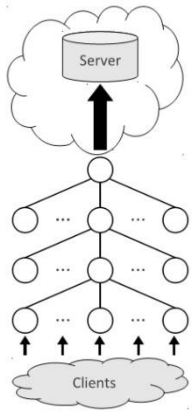

Fig. 2 Cache[13]

3.2 On-path caching

A most widely used cache practice in ICN. It deploys routers in the network for easy caching of Interest at PIT and data objects in various CS(s). The ability of the caches is thus studied as a predefined condition that enables adequate content availability when requests are posed. ICN promised advantages of the neighbor node producing results can therefore be achieved when more nodes cache contents based on the frequency of requests or recency of use (Least Frequency Use or Least Recently Use)[14-15]. For on-path caching, additional functionalities is less needed as the provision and intelligent of flooding contents is done with lesser complexity. Figure 2 presents a typical advantage procedure of caching in ICN. Figure 3 shows an easy cache practice on-path.

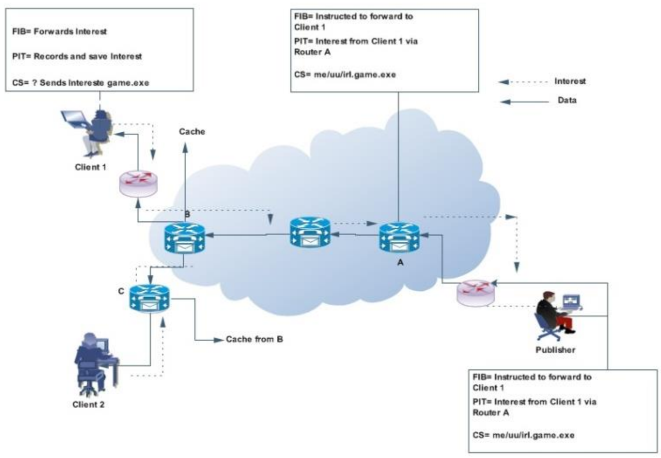

Fig 4. On-path caching

From Figure 2, the users referred to as clients, sends out the subscribed Interests off-path. Typically, this uses the PIT and FIB to deliver the Interest out of the network. Figure 3., appears with the advantage of gaining the flexibility of caching along network paths. This is seen as a benefitting advantage to cache hits, and availability of the content data. Several studies therefore aimed at carefully caching the contents, but still wary of the issue of cached-router out of network. This is thus seen as a challenge to ICN caching practice. B-COA aims at reducing this challenge through its built-in sections of waggling.

4. BEE-COLONY ALGORITHM

The Bee-colony study was introduced by Seeley in 1995 and extended by Seeley *et al.*,[16] as an optimization concept of treating a group interest. A study of the swarm nature of the bees was carried out by pointing out its advantages[17]. Further studies in the field of Intelligence, combinatorics and mathematics have seen the relevancy in its applicability. For the Bee, scouts and workers swarm out foraging. After the intensive duties of foraging, the waggle is then presented as a contest of advertisement.

For this study, B-COA will be proposed as an alternative of pollinating the network with a requested content data (data and content data are referred interchangeably). In ICN concept, when Interests are launched into the network as shown in Figure 3., the en-route Interests are tested to be guaranteed within the shortest possible time. However, the message as Interest is not having the identity of where to get the data. Studies in different approaches have thus found it worthy to catch the Interest as Pending Interest (PI) once it crosses a router that is

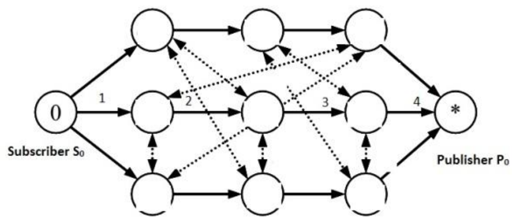

Fig. 3 Subscribing without swarm

unable to provide the data. So the PIT keeps this record

for subsequent similar Interest. However, once the needed data is found in the network through a publisher, the FIB charged with the function of returning the data as a route solver. Routers who failed the data delivery are thus followed and the record in the PIT is updated while caching takes place accordingly. The practice is thus similar to the foraging and waggling seen in the bee colony.

The predicted benefit of using the B-COA would be the uniform spreading of this needed data to improve the cache hit-ratio and mitigate delays during data searching.

5. PROBLEM FORMULATION

The ICN caching has widely been proposed to use the Leave Copy Down[1]. However, the contents are only cache-able to the routers that partook in the Interest finding. This is predicted to reduce the wide range of unattended routers that did not get to record the Interest request.

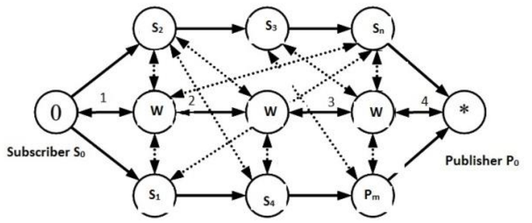

Fig. 5 Waggling added

B-COA is then applied to improve the delivery of the data on other neighbor routers through the swarm nature of the bee concept. In the Bee colony, nectar source address is populated by the waggling scouts. This yields to the swarm movement of other bees to the source. Inversely, the B-COA will be adopted as an inverse action to populate the cached data onto the network.

Suppose a network N of populated router R needs a large coverage of data Interest D.
Typically in ICN, the data is needed in almost all routers regardless of the path and route. This would improve data caching and fast cache hit.

Steps:

```
Subscribe data from a Publisher P0 at a
time t0
Cache on the routers with the PIT record
for subscriber S0 at t1
   Cached data D0 at the subscriber node R0
   should waggle after the   foraging to
   neighbors.
   Swarm by dispersing data D0 on all the
   neighbor routers R1 to Rn    within a
   given coverage
```

More specifically, Let Subscriber S = $\{S_0, S_1, S_2, ..., S_n\}$
Publisher P = $\{P_0, P_1, ..., P_m\}$,
Routeres R = $\{R_0, R_1, ..., R_n\}$,
with Data D = $\{D_0, D_1, ..., D_k\}$,
the above sets are all bounded by time t, and a cost of travelling and state (Frequency and Recency).
For an Interest to be sent, the following operation is needed
S_0 sents Interest I_0 via R to get Data D_0.

Furthermore from Figure 4, the subscriber initiates the Interest at action point 1, this request is traversed via routers on-path of the network until it gets the data from the publisher P_0.

This is seen as a form of data acquiring only by the routers linearly placed from the Figure 5. Our study yearn to cache contents based on the results obtained from the waggling dance calculation.
Waggling computation.
Data retrieved from the publisher at t_0 , is placed in the needed router.

Waggling W = The measure of Frequency of a data.
That is Recency C_i of a data > Frequency of usage the data F_i
such that:

When $C_i > F_i$, then swarm and propagate Data D_1 to all neighbors to cache (See Figure 5). In cache terms Most-Recently Used > Most Frequently Used. (MRU > MFU)

5.1 Proposed Bee- colony algorithm for ICN caching

```
1.  Set Network size
2.  Initialize the parameters
    (Subscribers, Data, Routers, Packets
    counts, Recency and Frequency values,
    count etc)
3.  While (Network size !=0)
4.      For each Interest (i) from
    Subscriber S; do
5.      Send Interest to first Router,
    then aquire random path
6.      If Interest is found; then update
    all routers R on-path
7.      Else change path
8.      Match Interest from Publisher P
    to the Subscriber S
9.      Return Data D
10.     If S(i) similar data at time ti ;
    then
11.     Test the MFU against MRU of the
    data by waggling W
12.     Swarm and propagate Data D  to
    all neighbors connected to R  with
    the value W to cache
13.     Return Data D
```

It can thus be concluded in four major steps as presented in Gallo et al., in [17] as:

```
for Interest Data D from a Subscriber S in
network; i = 1 to n,
for Data D retrieved from a Publisher P in
network; j = 1 to m,
```

```
            Forage;
                   Cache
                Waggle to test MFU and MRU
end for
end for
```


ACKNOWLEDGMENTS

This research is one of the Malaysian Government funded projects under the Ministry of Higher Education (MOHE) Long-Term Research Grant (LRGS) with reference number LRGS/TD/2011/UKM/ICT/02.

6. CONCLUSION

Attempts to propose the advantages in B-COA was presented in this study, even though it has not been implemented. The concept is seen as a promising idea to enhance the level and order of on-path caching in ICN. ICN idea is beneficial when host visitation number and counts are minimized through caching most recently used data packets. The paper concluded in theoritical composition to gain more Interests in formulating the algorithm in ICN terms. Future extensions will be the careful selection for the foraging and waggling scenarios of the Bee colony optimization algorithm.

REFERENCES

[1] V. Jacobson, D. Smetters, J. Thornton, M. Plass, N. Briggs and R. Braynard, "Networking Named Content," in *Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies*, 2009.

[2] I. Abdullahi, S. Arif and S. Hassan, "Ubiquitous Shift with Information Centric Network Caching Using Fog Computing," in *Computational Intelligence in Information Systems*, vol. 331, Springer International Publishing, 2015, pp. 327-335.

[3] G. Xylomenos, C. Ververidis, V. Siris, N. Fotiou, C. Tsilopoulos, X. Vasilakos, K. Katsaros and G. Polyzos, "A Survey of Information-Centric Networking Research," *Communications Surveys Tutorials, IEEE,* vol. 16, no. 2, pp. 1024-1049, Second 2014.

[4] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher and B. Ohlman, "A survey of information-centric networking," *Communications Magazine, IEEE,* vol. 50, no. 7, pp. 26-36, July 2012.

[5] I. Abdullahi, S. Hassan and S. Arif, "Prospective use of Bloom Filter and Muxing for Information Centric Network Caching," *ARPN Journal of Engineering and Applied Sciences,* vol. 10, no. 3, pp. 1169-1177, February 2015.

[6] B. Ahlgren, M. D'Ambrosio, M. Marchisio, I. Marsh, C. Dannewitz, B. Ohlman, K. Pentikousis, O. Strandberg, R. Rembarz and V. Vercellone, "Design Considerations for a Network of Information," in *Proceedings of the 2008 ACM CoNEXT Conference*, New York, NY, USA, 2008.

[7] C. Dannewitz, D. Kutscher, B. F. S. Ohlman, B. Ahlgren and H. Karl, "Network of Information (NetInf) - An Information-centric Networking Architecture," *Comput. Commun.,* vol. 36, no. 7, pp. 721-735, 2013.

[8] A. Dominguez, O. Novo, W. Wong and T. Valladares, "Publish/Subscribe communication mechanisms over PSIRP," in *Next Generation Web Services Practices (NWeSP), 2011 7th International Conference on*, 2011.

[9] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker and I. Stoica, "A Data-oriented (and Beyond) Network Architecture," *SIGCOMM Comput. Commun. Rev.,* vol. 37, no. 4, pp. 181-192, 2007.

[10] S. Wang, J. Bi and J. Wu, "Collaborative Caching Based on Hash-routing for Information-centric Networking," in *ACM SIGCOMM 2013 Conference on SIGCOMM, Hong Kong, China*, 2013.

[11] L. Wang, S. Bayhan and J. Kangasharju, "Optimal chunking and partial caching in information-centric networks," *Computer Communications*, 2014.

[12] G. Xylomenos, C. Ververidis, V. Siris, N. Fotiou, C. Tsilopoulos, X. Vasilakos, K. Katsaros and G. Polyzos, "A Survey of Information-Centric Networking Research," *Communications Surveys Tutorials, IEEE,* no. 99, pp. 1-26, 2013.

[13] M. Draxler and H. Karl, "Efficiency of On-Path and Off-Path Caching Strategies in Information Centric Networks," in *Green Computing and Communications (GreenCom), 2012 IEEE International Conference on*, 2012.

[14] C. Bernardini, T. Silverston and O. Festor, "Mpc: Popularity-based caching strategy for content centric networks," in *Communications (ICC), 2013 IEEE International Conference on*, 2013.

[15] W. K. Chai, D. He, I. Psaras and G. Pavlou, "Cache less for more in information-centric networks (extended version)," *Computer Communications* , vol. 36, no. 7, pp. 758-770, 2013.

[16] T. D. Seeley, P. K. Visscher and K. M. Passino, "{Group decision making in honey bee swarms}," *American scientist,* vol. 94, no. 3, p. 220, 2006.

[17] C. S. Chong, M. Low, A. Sivakumar and K. L. Gay, "A Bee Colony Optimization Algorithm to Job Shop Scheduling," in *Simulation Conference, 2006. WSC 06. Proceedings of the Winter*, 2006.

[18] I. Abdullahi, S. Arif, and S. Hassan, "Survey on caching approaches in information centric networking," *Journal of Network and Computer Applications*, vol. 56, pp. 48-59, 2015