An Analytical Model for Information Centric Internet of Things Networks in Opportunistic Scenarios

Jinze Yang, Yan Sun, Jesús Requena-Carrión, and Yue Cao, *Member, IEEE*

Abstract—The availability of environmental monitoring data collected by Internet of Things networks can be essential for many critical processes such as relief operations in disaster areas. The underlying communications infrastructure can be however severely compromised in these scenarios and therefore opportunistic approaches might be needed. Approaches based on Information Centric Networks (ICN), where moving devices forward collected data, have been proposed for opportunistic scenarios but to date, the dynamics of the delivery process in ICNs remain poorly understood. In this paper, we build a family of Markovian models for the delivery process of ICNs in opportunistic scenarios, that allows us to derive the end-to-end delay distribution and the storage ratio in terms of the encounter rate of the moving devices. Furthermore, we investigate how prefetching mechanisms affect the delivery process compared to conventional ICNs. The proposed models are fully validated in a computer simulation environment and demonstrate that the utility of delivery with prefetching reaches its peak in a short time and then decreases at a high rate. Our Markovian models can provide both the insight and quantitative estimations that are needed to design practical ICNs in opportunistic scenarios.

Index Terms—Information Centric Networks, Opportunistic Scenarios, Prefetching, Markovian Model

I. Introduction

Internet of Things (IoT) technologies are essential for automatic data collection in remote and post-disaster scenarios. In [1], an IoT based Delay Tolerant Network (DTN) scheme was proposed to deliver information to residents in remote areas without communications infrastructure and under extreme weather conditions. In [2], an IoT network was proposed to assess damages caused by disasters. By deploying sensor nodes around a building, its tilt and shift could be measured, and based on them the damage could be evaluated. However, although environmental data are essential in remote and disaster areas, the delivery of data can be extremely challenging due to a deficient communications infrastructure.

To address the challenges of the lack of a working communications infrastructure, previous studies have treated such scenarios as a subset of opportunistic scenarios involving moving devices. In such scenarios, end-to-end routes exist in very rare occasions and thus, a successful data delivery relies on

(Corresponding author: Jesús Requena-Carrión)

J. Yang, Y. Sun and J. Requena-Carrión are with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mail: jinze.yang@qmul.ac.uk; yan.sun@qmul.ac.uk; j.requena@qmul.ac.uk)

Y. Cao is with the School of Computing and Communications, Lancaster University, Lancaster LA1 4WA, U.K. (e-mail: yue.cao@lancaster.ac.uk)

the interplay of data-carrying moving devices that eventually encounter the recipient device. In such scenarios, DTNs and Information Centric Networks (ICNs) are among the most popular solutions. Delay allowances allow devices to store data for longer periods, thus increasing the encounter probability between data-carrying devices and destination devices. As in conventional DTNs, destination devices are identified by unique host addresses, for instance, by endpoint identifiers [3]. Due to the address-oriented delivery, the data-carrying devices cannot deliver the data to any newly joined destination device yet appearing in the destination list once generated [4].

Following content-oriented delivery design principles, ICNs have been introduced to work alongside DTNs in opportunistic scenarios [5]–[8]. In ICNs formulations moving nodes carry data units, which are commonly referred to as content and have a unique identifier. Destination nodes, also known as consumers, obtain the required content from data-carrying nodes they interact with. Prefetching mechanisms, which produce expected contents in advance, are introduced into ICNs to further reduce the end-to-end delay [9], [10].

In opportunistic scenarios, the end-to-end delay is a performance factors hard to guarantee. Therefore, the probability of successful delivery within delay allowance is commonly the main contributor to the Quality of Service (QoS) assured services, where the term "delay allowance" corresponds to the maximum allowed delay for content delivery. Amongst the previous studies dealing with opportunistic scenarios, only a few of them focus on exploring the analytical relationship between delivery probability and delay allowance. The authors in [11] studied the delivery process in DTNs and proposed two analytical models. Both analytical models were developed for source to destination one-way transmission and are thus unfit for ICNs in opportunistic scenarios where both request sent and content delivery should be taken into consideration. In [12] a method to obtain the delivery rate with a certain content number was proposed. Since the method makes use of a fixed delay allowance and content number to produce the result, the relationship between delivery and delay is yet to be discussed. Moreover, due to the conflicts between the limited content storage at the nodes and the long delay allowance during content dissemination, the mathematical analysis of two factors is desirable. With prefetching mechanism, storage consumption can be further improved.

Aiming to study the quantitative relationship between delivery probability and delay in DTN-ICN, in conjunction with prefeteching, two proposed models in this work for

analysis are named as Delay distribution of CONventional ICN (CONI-D) [13] and Delay distribution of ICN with All Agents Prefetching (IAAP-D) [1]. The Storage ratio models for CONcentional ICN (CONI-St) and ICN with All Agent Prefetching (IAAP-St) are defined as well to describe the percentage of nodes having content versus delay allowance. Finally, a utility function is established to study the storage ratio and delay distribution jointly. The main contributions are listed below:

- Two storage ratio models, CONI-St and IAAP-St are proposed to reflect how the content dissemination changes through time. Both CONI-St and IAAP-St can produce accurate estimation of storage along with time evolution for the given network configuration. The IAAP-D is also proposed to model the delay distribution of IAAP given elapsed time, comparing with CONI-D [13].
- A novel emulation method is designed to analyse the concurrence of IAAP and CONI. The proposed method can produce a good estimation of the prefetched content number in IAAP.
- The utility function of delivery ratio is defined to illustrate the trade-off between delay reduction and storage usage. The optimal request lifetime is then to be obtained from the utility function.

This paper is organized as follows: In Section II, recent works are reviewed. In Section III, the models for CONI and IAAP are introduced, and the distribution of end-to-end delay and storage ratio are obtained. In Section VII, the model is fully evaluated via simulation in the scenario of an ICN with nodes following Random Waypoint Mobility Model (RWMM). Then, Section IX contains the conclusions of our study. Finally, in Section VIII, the unsolved challenges are discussed.

II. RELATED WORK

Since the IoT in opportunistic scenarios relies on mobility aided DTNs, the mobility models and mobility based DTN analytic models are reviewed first. Then, to study the IoT combined with ICNs and prefetching, the DTNs based ICNs protocols are discussed. Finally, the prefetching related works are reviewed.

A. Mobility Models

Screening the latest works, a noticeable percentage of forwarding protocols adopted in IoT networks are opportunistic based. Therefore, the data forwarding is highly relying on the node movements. Naturally, the analytical models are built on top of mobility models. The mobility models can be classified into two categories: the real world traces and the model based trajectories. The real-world traces [14] are realistic but fixed once finished, result in little scalability and probable specialized optimization. On the contrary, the model-based trajectories, which might not reflect exact human movement, are much more flexible and scalable.

The mobility models, such as Random Waypoint Mobility Model (RWMM) [15] and Random Walk Model [16], are well studied in many research works. The RWMM works as that a moving node randomly chooses a destination point within an area. When the node reaches its destination, it will stay for a random length of time and subsequently chose a new destination for the next movement. Since RWMM is known as a trustworthy model for human walking in a certain area, such as in villages, many research works have adopted RWMM for evaluating protocols' performance in DTNs and Mobile Ad hoc NETworks (MANETs), bearing with its limitations [17]. Moreover, many simulation platforms, such as the Riverbed Modeler (formal OPNET) [18] and the ONE [19], integrate RWMM to support research works. In favour of the popularity, the RWMM is adopted in this paper. According to [20], many mobility models and traces have the property that the encounters of two nodes follow the Poisson distribution. Thus, these mobility models can be modelled as a series of encounters. Since the proposed models take the Poisson distribution as an input, it provides the advantage that the proposed model can be applied to any mobility models following the Poisson distribution.

B. Delay Tolerant Networks

Along with the mobility models, the Markov chain has turned out to be a common method to analyse the IoT network protocols. According to different types of wireless connectivity, IoT networks can be classified into two categories: MANETs and DTNs. The MANET assumes that usually there are end-to-end routes, while DTNs assume that there are rarely end-to-end routes. Since this work focuses on opportunistic scenarios, only DTN is discussed in this paper. The Markov chain is adopted in [11] which studies both Epidemic routing protocol [21] and two-hop multi-copy protocol. The Epidemic protocol [21] allows all nodes to send contents to every encountered node, while the two-hop multi-copy protocol only allows content producer to send contents. The authors of [11] take the increasing number of content as state transitions, and use the mean encounter interval to calculate the mean end-to-end delay. However, the mean encounter interval is different between each state, thus the mean end-to-end delay conducted by [11] would not be very accurate. Diana et al. modelled the binary Spray and Wait (SnW) protocol [22] with a complicated Markov chain. The SnW protocol sets maximum number of content copies in advance. Although SnW allows every node to forward the content copies, the copies will not over consume much resources due to limited maximum number of copies. Besides, Diana et al. and other researchers extended the Markovian model to more realistic cases where mobile nodes have heterogeneous mobility models [23], [24].

It is worth pointing out that both [11] and [24] only produce a mean end-to-end delay without further investigating on the influences of other factors, such as content lifetime, etc. Moreover, only one-way transmission is modelled in [11], [23], [24] to reflect that one node sends a content to another node. When considering ICNs, data transmission normally composes of both requesting and responding phases, which has been overlooked by above mentioned works.

---

[1]The agent here refers to the nodes which can forward contents on behalf of either consumer or producer.

C. Information Centric Networks

In ICNs, a consumer can request one specific content with one specific request, regardless of the content producer. Decoupling of content and producer makes the forwarding flexible and increases the delivery probability. Therefore, works combining ICNs and DTNs have attracted quite attentions. Recently, researchers published a series of works to investigate the ICNs performance in opportunistic scenarios. General advance of ICNs over traditional DTNs has been studied in [4]. The performance factors of [4] are also adopted in this paper. In [12], a content dissemination scheme similar to Content Centric Networking (CCN) [8] is adopted. Content consumers first broadcast requests for the desired contents. Upon receiving a request to content, content producers or content caching nodes response to the consumer with contents and eliminate the requests once fulfilled. However, the mobility of content producers further introduces a new challenge as when producers are moving, it would be difficult to find and fetch new contents. In [25], [26], the described scenario is an IoT network where hikers relay information. Anastasiades et al. paid special attention to the producer mobility management issue which has greater impact on performance than consumer mobility. In [26], a content discovery approach is proposed to resolve the content name when a mobile producer is discovered. However, the proposed content discovery approach has convergence issue in opportunistic scenarios. Although the evaluations are mainly carried out by testing equipments, only a one-dimensional mobility model is mentioned in [25]. In work [27], authors developed a routing protocol based on social ties and evaluated with the random walk mobility model. Aforementioned works did not exploit much factors which affect the whole delivery process.

In our previous work [13], we exploited the relationship between end-to-end delay and encounter numbers regardless of storage consumed by contents. Concluded from [13], the end-to-end delay is determined by mobility models, total number of nodes and initial number of contents in the network. However, in [13], the storage, which is another limited resource in disaster scenarios, is not taken into consideration.

D. Prefetching

In order to reduce the latency of fetching contents, the prefetching have been studied for years. The proactive caching means caching before contents are requested while the caching means storing contents in local storage which is different from prefetching. In this paper, the term proactive caching and prefetching are interchangeable. Different frameworks and algorithms are proposed for ICNs [9]. In [28], a prefetching mechanism is introduced based on predictions of vehicles' mobility and predictions of popular contents. The popular contents are prefetched to cache nodes where large traffic occurs. The popular contents are predicted by Prediction-by-Partial Matching (PPM) algorithm. In PPM, a Markov chain is built according to local history. When the request sequence matches one of the chains, the predictions are made [29]. Similarly, the [30] also developed a mobility prediction based prefetching approach. Both [28] and [30] provide a solution
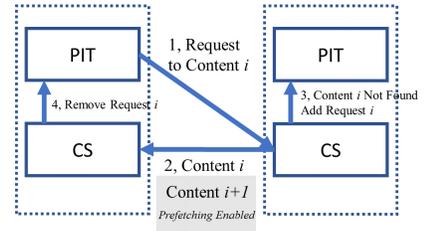

Fig. 1. Content Forwarding Data Structure

for which node should prefetch contents. However, most of the similar works focus on vehicular networks which rely on fixed road side units and wired networks. Few works consider opportunistic scenarios. The [31] proposed a fuzzy path model based mobility prediction approach for data prefetching for wild animal monitoring which is one of the opportunistic scenarios, but ICNs are not adopted.

The above-mentioned works optimize storage consumption via selecting prefetching devices based on their locations. In opportunistic ICN scenarios, it is rare to have static devices helping with the content delivery. Moreover, the forwarding capability of static devices is overlooked in most of the networks. Therefore, the performance of mobility prediction based prefetching could be affected. Meanwhile, prefetching can increase the storage consumption dramatically. Thus, in this work, we focus on the network level storage increase and end-to-end delay decrease when modelling prefetching in ICNs.

E. Summary

The above literature review suggests that network-level analytical model of mobile ICNs is still a rarely touched field. Researchers are still exploring different mechanisms. By modelling the storage consumption and delay distribution, a complete view of how factors affect the delivery is provided. Meanwhile, research works can be guided and accelerated with the proposed analytical models. In the next section, the system model is described.

III. SYSTEM MODEL

A. Scenario definition

The post-disaster scenario is chosen to study the content dissemination of ICN in opportunistic scenarios. The target scenario is a square region with side length a. In the region, M disaster-relief staff, who follow the RWMM with velocity v, search in the area with devices (nodes), such as smart phones. The nodes in the region but incompatible with ICNs are not considered. Node failure rate is assumed extremely low and ignored in this work. The maximum distance, which allows instant requests and contents transmission affected by noise or obstacles, is denoted as r. All these parameters are combined to determine a pair-wise encounter rate λ_p which will be discussed in following sections. As mentioned in Section II-A, the proposed model can be extended to any mobility models

and traces in which encounters between nodes follow the Poisson distribution.

The M nodes can be classified into three types: producers, consumers and agents. The producer has all the contents from the beginning of simulation, for instance, the producer is a device which can connect to the Internet via satellite networks. Meanwhile, the content lifetime is t_c, and the content generation time is set as the time point when the content is sent out by producer the first-time. The consumer requests contents with request generation rate R and request lifetime t_r. The agents can request contents on behalf of consumers. Upon receiving contents, agents can also disseminate contents on behalf of producers. However, they cannot generate either requests or contents. Each agent and each consumer can hold maximum B content units. The requests and contents will be discarded when they expire or the buffer is full. The First In First Out (FIFO) is implemented as buffer replacement algorithm. The following explanation assumes that there is one producer, one consumer and $M - 2$ agents in this scenario. The proposed model can be extend to embrace multi producers and consumers cases by reconfiguring the initial setup.

B. Description of CONI

The CONI is the analytical model which models the Epidemic-type forwarding protocol [12] with FIFO for CCN [8]. The Epidemic-type CCN forwarding protocol [12] is a distributed forwarding method, and adapts to the DTN with minimum changes, which reflects the performance of ICNs in opportunistic scenarios without prefetching. Thus, we develop CONI to study the delivery process of ICNs in opportunistic scenarios. The CONI-D is proposed by us in [13] while CONI-St is proposed in this paper.

The forwarding protocol is described in this paragraph. There are two data structures helping completing the delivery in every node's forwarding layer. The Pending Interest Table (PIT) records every unfulfilled request which are exchanged during every encounter, while the Content Storage (CS) stores every received content. When two nodes encounter each other, the PIT is exchanged as one packet, and demonstrated as action 1 in the Fig.1. Received PIT is checked according to CS, and matched contents are sent as responses, shown as action 2 in the Fig.1. Unfulfilled requests are stored in PIT, shown as action 3 in the Fig.1. Upon receiving a new content, the fulfilled request is removed, demonstrated as action 4 in the Fig.1. The name resolution is out of the scope of this paper.

C. Description of IAAP

According to our review in Section II, since the prefetching mechanism can further reduce end-to-end delay by consuming more storage, we develop IAAP to investigate how much performance difference prefetching can achieve and how the prefetching affects the following content dissemination. For IAAP, the producer is assumed to be capable of predicting the next most popular content, denoted by content $i + 1$, based on received requests. Since the prediction hit ratio and the predicted storage ratio have the linear relationship, to allow the focus of the model analysis, it is assumed that the hit ratio

of prediction is 1 in this paper. The loosened-assumption cases are discussed in Section VIII. The forwarding mechanism in IAAP is different from CONI: once a node receives the request to content i, if the node has both content i and content $i + 1$, the node replies to the request sender by both contents and the prediction results, which is demonstrated as action in shadow in the Fig.1. Furthermore, all agents in IAAP are capable of prefetching. Thus, this mechanism is called All Agents Prefetching (AAP).

D. Performance metric

The evaluation parameters chosen in this paper is the end-to-end delay distribution and storage ratio. The delay distribution is introduced which describes the delivery probability versus delay allowance. Since in opportunistic scenarios, given an infinite delay allowance, the consumer will eventually encounter the producer or a cache node. Thus the delivery probability will be 1 with infinite delay allowance. However, since the resources are usually limited in opportunistic scenarios, it is impractical to allow contents and requests to be valid for unlimited time. The delay distribution can provide delivery probability with certain delay allowance. By deriving the delay distribution, the delivery probability changing rate can be obtained, which can help determine a effective delay without much delivery probability lost. The storage space, as another crucial resources, is also important for data transmission. When delay allowance increases, the contents will be disseminated to more nodes in the networks. Therefore, the storage ratio model is built to describe the number of content copies disseminated in the networks. Based on the proposed models, the impact of varies factors are quantified while the relationship between delivery probability and storage ratio is addressed. Other evaluation parameters, such as CPU and overhead, are less relevant to forwarding. The forwarding mechanism is similar to CCN which is simple, thus CPU requirement is low. Meanwhile, the calculation of proposed model is carried out during network planning stage, thus no extra overhead or CPU power is required.

Finally, from the network aspect, given a content i, let the number of nodes with request to content i be O_i and the number of nodes with content i be N_i. Therefore, the sum of the O_i and the N_i is smaller or equal to total number of nodes M. Above condition is presented by $O_i + N_i \leq M$. The proposed notation is summarized in Table I.

E. Assumption

Since the transmission happens when one node encounters another, the delivery of contents can be represented by a series of encounters. During every encounter, the probability of transmission can be represented by the encounter probability of nodes with requests or contents. The impact of transmission can be represented by the transition between network states (O_i, N_i). Therefore, the delivery process can be modelled by absorbing finite state Markov chain [11]. Before defining the models, the following assumptions are made:

- The communication range r is much smaller than the side length of the region a, $r/a << 1$, which implies

M	Total number of nodes
a	Length of area (m)
v	Host speed (m/s)
r	Communication range (m)
R	Request rate (1/s)
B	Agent buffer size (content unit)
t_c	Content lifetime (s)
t_r	Request lifetime (s)
t	Time (s)
n	Encounter number
N_i	Number of nodes with content i
O_i	Number of nodes with request to content i
λ	Encounter rate of one node (1/s)
λ_p	Encounter rate of one pair of nodes (1/s)
λ_M	Encounter rate of all nodes (1/s)

that encounters of three or more nodes simultaneously are very unlikely.

- It is assumed that both request lifetime t_r and content lifetime t_c are long enough to complete the delivery process. Thus, both O_i and N_i can only be changed during an encounter.
- To guarantee the independence of each delivery process, requests generated within the interval between two consecutive encounters are considered as one single request. This is because the delivery process of these requests is exactly the same for each request, indicating that these deliveries are dependent to each other.
- The agents have large enough buffer size B. Therefore, the O_i and the N_i will not be changed because of a full buffer.

IV. DERIVATION OF STORAGE RATIO MODEL

In this section, the calculation of storage ratio is introduced first, which takes the following CONI-St and IAAP-St as inputs, and produces quantitative results. The Markovian models of CONI-St and IAAP-St are described next, which reflect the encounter-based location-independent content dissemination mechanisms. Next, since the IAAP-St is likely to work with CONI-St, the emulation method is introduced to estimate the percentage of contents disseminated with each model. Finally, the performance differences between the CONI and IAAP are discussed.

A. Storage ratio calculation

The storage ratio is defined as a percentage of nodes with content in this paper. The producer is excluded here, because the producer always has contents. The storage ratio $f_{stoR}[n]$ can be represented by:

$$f_{stoR}[n] = \frac{f_{sto}[n] - 1}{M - 1} \qquad (1)$$

where the f_{sto} is the expected number of nodes with content after n encounters. The f_{sto} is defined as:

$$f_{sto}[n] = CZ_{sto}^{(n)} \qquad (2)$$

where C is a 1-by-x vector describing the N_i, which associates with every state of state distribution vector Z_{sto}. The Z_{sto} is a x-by-1 vector which describes the probability of the network's current state. The network state here, as introduced in the previous section, is denoted by (O_i, N_i). Let the initial state distribution be Z_{sto0}. The probability of reaching each state after n times transition, denoted by $Z_{sto}^{(n)}$, can be obtained by:

$$Z_{sto}^{(n)} = Q_{sto}^{(n)} Z_{sto0} \qquad (3)$$

where the x-by-x transition matrix Q_{sto} is defined as the probability of transition from one state to another. Details about how to obtain transition matrix Q_{sto} are introduced in following sections.

B. Storage ratio model of CONI

Based on the above-mentioned assumptions, the transit state of the CONI-St is denoted by (O_i, N_i) where the number of nodes with request to content i must be greater than 0. The above condition is presented by: $O_i > 0$. The absorbing state, which always transits to itself, is defined as $(0, N_i)$ which means that all the requests are satisfied. The described delivery process can be summarized into two basic actions: replication of requests and replication of contents. Therefore, there are three types of events during an encounter: Request Increase Transition (RIT), Content Increase Transition (CIT) and Self Transition (ST). The transition probability can be obtained by dividing the number of node pairs which will trigger the type of transition by the total number of node pairs. For instance, when there is 1 consumer, 1 producer and 3 agents with no requests or contents, the total number of node pairs is 10. The probability of CIT happens during 1 encounter is $1/10$ where the 1 is the only pair of nodes, namely consumer and producer. Given a state (O_i, N_i), these three types of events are shown in Fig.2 and explained in detail:

- RIT: one node with request to content i encounters a node and replicates the request to the encountered node. The state transits to $(O_i + 1, N_i)$ and the probability of RIT occurring is $P\{RIT\} = 2O_i(M - O_i - N_i)/M(M - 1)$. An example is demonstrated in Fig.2a.
- CIT: one node with content i encounters a node with request to content i and replicates the content i to the encountered node. The current state transits to $(O_i - 1, N_i + 1)$ and the probability of CIT occurring is $P\{CIT\} = 2O_i N_i/M(M - 1)$. When O_i reaches 0, the absorbing states are reached. An example is demonstrated in Fig.2b.
- ST: all the other cases of encounters belong to this case. The target state is (O_i, N_i). The probability that this kind of transition occurs when two nodes encounter is $P\{ST\} = 1 - P\{RIT\} - P\{CIT\}$. An example is demonstrated in Fig.2c.

As an example, a 5-node CONI-St Markovian state transition diagram is shown in Fig.3a. The transition probabilities of CONI-St can be represented by a probability transition matrix
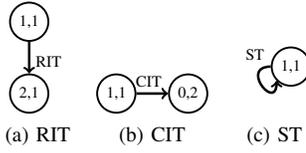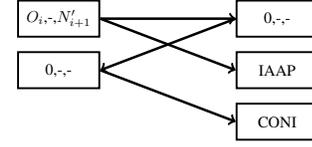
Fig. 2. Different types of transition

Fig. 3. 5 nodes Markovian state transition diagram

Fig. 5. Transition diagram of different initial states of IAAP

Q_{sto}. Due to limited pages, the following 5-by-5 probability transition matrix, which describes the transition probability between each states of a 3 nodes CONI-St model, is shown instead of the 14-by-14 transition probability matrix of the 5 nodes CONI-St model:

$$Q_{sto3} = \begin{bmatrix} 0.33 & 0.33 & 0 & 0.33 & 0 \\ 0 & 0.33 & 0.67 & 0 & 0 \\ 0 & 0 & 0.33 & 0 & 0.67 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

With the proposed (1), the storage ratio can be obtained. To convert from encounter number scale to time scale, the time point is obtained by encounter number n divided by λ_p. The above condition is presented by $t = n/\lambda_p$. The λ_p is the encounter rate of any pairs of nodes. To demonstrate how the delivery consumes the storage, the CONI-St with different M is shown in Fig.4a. The increase of λ_p, caused by increase of M, results in faster storage ratio increase and higher upper bound.
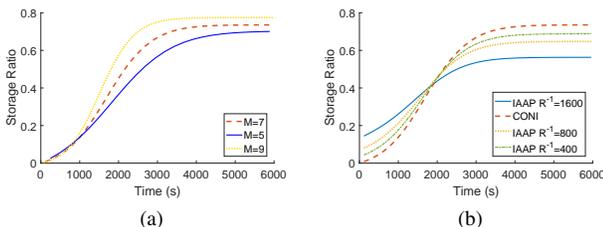
Fig. 4. (a) Time evolution of CONI-St with different total number of nodes M, (b) Time evolution of 7 nodes IAAP-St with different request generation rate R

C. Storage ratio model of IAAP

As introduced in Section III, the prefetching scheme, which is studied in this paper, is also based on Epidemic protocol [21]. To simplify the notation, the predicted next most popular content is denoted by content $i + 1$. When a node receives request to content i, it checks both content i and content $i+1$. Other forwarding mechanism remains the same. Because all the contents are initially stored in producer, the prefetched contents $i + 1$ are always bound with content i. Therefore, there are two possible contents, namely, content i and joint content i and $i + 1$, can satisfy request to content i. The joint content i and $i + 1$ is denoted by joint content $i + 1$.

Because of the newly introduced joint content $i+1$, the state (O_i, N_i) used in the previous section can no longer represent the state of networks. Therefore, triplets (O_i, N_i', N_{i+1}') are used here to represent the more complex network state. Here, the N_i' denotes the number of nodes with only content i, and the N_{i+1}' denotes the number of nodes with joint content $i+1$. Thus, the N_i is the sum of N_i' and N_{i+1}'. The above-mentioned equation is presented by: $N_i = N_i' + N_{i+1}'$.

Unlike CONI-St, the IAAP-St has four types of events which are listed below:

- RIT: when a node with request to content i encounters a node and replicates the request to the encountered node. The target state is $(O_i + 1, N_i', N_{i+1}')$ and the probability of transition is $P\{RIT\} = 2O_i(M - O_i - N_i)/M(M - 1)$.
- CIT: a node with request to content i encounters a node with only content i. The content i is replicated to the node. The target state is $(O_i - 1, N_i' + 1, n_{i+1}')$ and the probability of transition is $P\{CIT\} = 2(O_i - 1)N_i'/M(M - 1)$. When O_i reaches 0, the current state transits to absorbing states.
- Joint content Increase Transition (JIT): similar to CIT, the target state is $(O_i - 1, N_i', N_{i+1}' + 1)$ and the probability of transition is $P\{BIT\} = 2(O_i - 1)N_{i+1}'/M(M - 1)$.
- ST: all the other cases belong to this case.

The Markovian state transition diagram of IAAP-St is shown in Fig.3b. To simplify the diagram, the ST is omitted.

Obviously, joint content $i + 1$ will affect the initial state distribution of request to content $i+1$. The prefetched contents in the networks break the independence of dissemination of consecutive contents. In the next subsection, this issue is addressed.

D. Dependency between two consequent requests

In the real world, it is likely that there is more than one request transmitted in the networks at the same time. When the AAP is adopted, one request can trigger the transmission

of two contents which could lead to a very complicated dependency of the following requests. However, according to our investigation via Riverbed Modeler [18] simulation, the matched requests still have a major impact of contents transmission. Details of Riverbed Modeler [18] simulation are discussed in Section VII. To better understand the statistic property behind the mechanism, a simplified model is discussed first and verified under complicate simulation settings later. Before defining the model, the following assumptions are made:

- The requests are sent periodically, and the contents $i+1$ are available to be prefetched during this period. In other word, only one request affects the content $i+1$ in the network at one time.

Based on the above assumptions, the content transmission can be modelled by three initial cases. First, the consumer has content i before it generates request to content i. This initial case will lead to an instant content delivery. Therefore, none of content $i+1$ will be prefetched. This case is denoted by $(0, -, -)$. As a consequence of of $(0, -, -)$, when response to request to content i, the copies of content i are always bound by contents $i+1$. Thus, CONI-St is chosen to model the second initial case. When content $i-1$ is prefetched by agents, there are both contents i and joint contents $i+1$ in the network. The IAAP-St should be chosen to model the third initial case. Based on these three initial cases, the final cases of previous transmission can be summarized as $(O_i, -, N'_{i+1})$ and $(0, -, -)$, which means consumer has or has not send request to content i accordingly. As shown in Fig.5, once the copies of request to content i are sent out, the initial cases can be either $(0, -, -)$ or IAAP-St. Otherwise, the transmission starts from CONI-St.

According to the dependency between two consecutive requests, an adapting tool can be produced which maps the final state distribution of content i to initial state distribution of content $i+1$. Therefore, both IAAP-St and CONI-St are used together to model the content transmission of IAAP. The stationary distribution of initial cases is denoted by $[p_{0--}, p_{IAAP}, p_{CONI}]$, where each item is the probability of the transition start from $(0, -, -)$, $IAAP$ and $CONI$ accordingly. The stationary distribution can be reached within several content transmissions, and is used to calculate the storage ratio and delay distribution.

Similarly, the expected number of nodes with content of IAAP can be obtained by the weighted sum of expected number of nodes with content f_{sto} of three different initial cases. The weight is $[p_{0--}, p_{IAAP}, p_{CONI}]$. The f_{sto} with different request rate R is shown in Fig.4b. When R increases, the available time for prefetching increases, thus the prefetched storage ratio, which indicates the storage ratio at time 0, increases. With higher prefeteched storage ratio, the upper limit of storage ratio becomes lower because the requests are more likely to be consumed than CONI. However, in real networks, because the request rate R might not follow any distribution, it is difficult to stop prefetching at certain pre-set time. Thus, this model mainly focuses on the estimation of prefetched storage ratio and how the prefetching impacts the
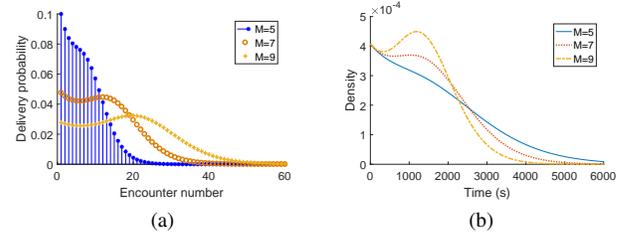

Fig. 6. (a) PMF of encounter number with different total number of nodes M, (b) CONI-Delay distribution with different total number of nodes M

delay distribution.

V. DERIVATION OF DELAY MODEL

In this section, similar to storage ratio model, the calculation of delay distribution is introduced first. The Markovian model of IAAP-St is introduced, which are input of delay distribution calculation. Then, the stationary distribution of initial cases of storage ratio is borrowed to produce the delay distribution of IAAP.

A. Calculation of delay distribution

Similar to the state transition matrix of storage ratio model, the transition probabilities of CONI-D [13] and IAAP-D can be represented by a probability transition matrix Q, which is introduced in the following section. Meanwhile, the current network state distribution can be described by a x-by-1 vector Z. Let the initial state distribution be Z_0, the probability of reaching each state after n transition $Z^{(n)}$ can be obtained by:

$$Z^{(n)} = Q^{(n)} Z_0 \tag{4}$$

Let $Z_A^{(n)}$ denote the probability that the network is in absorbing state $(0, -, -)$ after n transitions. The Probability Mass Function (PMF) $S_A(n)$, which describes the probability that the network reaches $(0, -, -)$ with exact n transitions, is defined by:

$$S_A(n) = \begin{cases} Z_A^{(1)} & \text{if } n = 1 \\ Z_A^{(n)} - Z_A^{(n-1)} & \text{otherwise} \end{cases} \tag{5}$$

The $S_A(n)$ with different total number of nodes M is illustrated in Fig.6a. When the total number of nodes M increases, the required encounter number for delivery also increases. Therefore, the delivery might require long delay. However, the total number of nodes M changes the λ_p at the same time. It is difficult to conclude the impact of total number of nodes M here. Moreover, when the total number of nodes M increases, the number of agents with neither contents nor requests is higher at beginning. Therefore, the encounters without requests or contents transmission are more likely to happen than any other transitions. More agents result in low deliver probability initially.

In order to obtain the delay distribution, it is necessary to find out the distribution of the time between consecutive encounters, which can convert the delivery probability versus encounter to delivery probability versus time.

The interval t between two consecutive encounters in RWMM follows the exponential distribution [20], which can be described as:

$$P\{T_j < t\} = 1 - e^{-\lambda t} \qquad (6)$$

where T_j is the encounter interval of node j. The parameter λ is the encounter rate of one node encountering any other nodes, and λ is greater than 0 ($\lambda > 0$). The t is observation time. This property is examined and used in [12], [20], [32]. Nevertheless, the encounters of node j are composed of encounters of $M-1$ pairs of nodes, each happens identically and independently. Thus, based on infinite divisibility of exponential distribution, the encounter of a random picked pair of node j and node h follows:

$$P\{T_{jh} < t\} = 1 - e^{-\lambda_p t} \qquad (7)$$

where λ_p is the encounter rate of one pair of nodes and can be obtained by: $\lambda_p = \lambda/(M-1)$. Similarly, based on the number of unique pairs of nodes, the distribution of encounters of any pair of nodes follows:

$$P\{T_M < t\} = 1 - e^{-\lambda_M t} \qquad (8)$$

where λ_M is the encounter rate of any pair of nodes and can be obtained by: $\lambda_M = \lambda_p M(M-1)/2$. Since the time interval between two consecutive encounters of any pair of nodes follows the same exponential distribution with parameter λ_M, the delay of delivery with n encounters corresponds to sum of n exponential distribution, which is a gamma distribution $\Gamma(k, 1/\lambda_M)$ with scale parameter n and shape parameter $1/\lambda_M$.

By combining the delivery probability with n encounters and gamma distribution with scale parameter n, the PDF of delivery probability vs. delay is described by (9):

$$f_{mix}(x) = \sum_{n=1}^{\infty} S_A(n) * \Gamma(x; n, 1/\lambda_M) \qquad (9)$$

where $S_A(n)$ is the probability of delivering a content with n encounters. The $f_{mix}(x)$ with different total number of nodes M is shown in Fig.6b. The CONI-D [13] can be interpreted as delivery probability within certain time, which can be interpreted as a certain percentage of content can be delivered within certain delay allowance or request lifetime t_r. It is worth noting that the delivery probability with higher M is higher than the delivery probability with lower M which is different from the Fig.6a. Although the increase of encounter rate λ_M reduces the delivery probability per encounter, overall the delay is reduced by the shorter encounter interval.

B. Delay model of IAAP

Similar to storage ratio models, the transit state of IAAP-D is denoted by $(1, O_i - 1, N_i)$, each item means the number of consumer, the number of agents with request to content i and the number of nodes with content i accordingly. The absorbing state, which always transits to itself, is defined as $(0, -, -)$, which means that the consumer gets the content i regardless of the number of others. The described delivery process can be summarized into three basic actions: replication of requests,
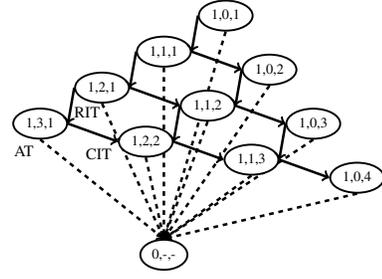

Fig. 7. Transition diagram for 5 nodes IAAP-D

replication of contents, and content delivery. Therefore, there are four kinds of events which may happen during encounters: Absorbing Transition (AT), RIT, CIT and ST. Given a state $(1, O_i, N_i)$, the four types of events are explained in detail:

- AT: the consumer encounters a node with content i. The content is delivered to the consumer and the state transits to $(0, -, -)$. The probability of AT happening is $P\{AT\} = O_i/M$.
- RIT: one node with request to content i encounters a node and replicates the request to it. The state transits to $(1, O_i + 1, N_i)$ and the probability of RIT occurring is $P\{RIT\} = O_i(M - O_i - N_i)/M$.
- CIT: one node with content i encounters a node with request to content i and replicates the content to the encountered node. The state transits to $(1, O_i - 1, N_i + 1)$, and the probability of CIT occurring is $P\{CIT\} = (O_i - 1)N_i/M$. Since the contents are only transmitted upon requested, CIT only occurs when the number of nodes with request to content i is greater than 1. The above condition is presented by $O_i > 1$.
- ST: all the rest of cases of encounter belong to this case. The state transits to itself. The probability of ST occurring is $P\{ST\} = (M - O_i(M - O_i))/M$.

The transition diagram of IAAP-D for 5 nodes is demonstrated in Fig.7.

Recall the assumptions made in Section IV-C. After request interval R^{-1}, only content i can satisfy request to content i, thus the state transition model of IAAP-D is the similar to the state transition model of CONI-D [13]. With the convergence distribution $[p_{0--}, p_{IAAP}, p_{CONI}]$, initial distribution of IAAP-D can be obtained. The PDF of 7 nodes IAAP-D with different request rate R is shown in Fig.8a. With higher prefetched storage ratio, the delivery probability within 500s is quit high compare to CONI-D [13]. The delivery probability of smaller request rate R is higher than the delivery probability of the larger one within 500s.

VI. UTILITY FUNCTION

Since the proposed models is capable of estimating performance with the knowledge of factors, the cost effective network planning can be achieved in practice. In this section, a utility function, which describes the relationship between delivery probability gain and storage consumption with regards
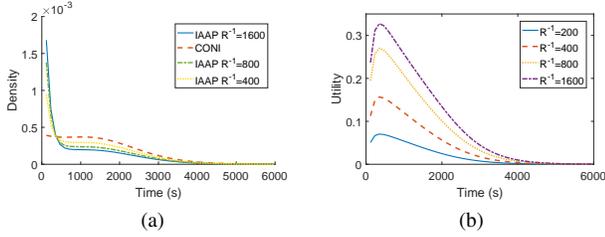
Fig. 8. (a) 7 nodes IAAP-Delay distribution with different request generation rate R, (b) Utility curve of $M = 7$

Fig. 9. Analytical time evolution of CONI-Storage ratio vs. simulated CONI-Storage ratio, $t_r = 6000s$, $t_c = 6000s$, $R^{-1} = 100s$

to certain t_r, is defined to find an optimal t_r. The utility function is defined as:

$$f_{ut}(x) = (1 - f_{cost}(\lceil x\lambda_M \rceil))f_{gain}(x) \qquad (10)$$

where $f_{cost}(\lceil x\lambda_M \rceil)$ is the cost function, which describes the extra storage consumption caused by prefetching mechanism. Meanwhile, $f_{gain}(x)$ is the gain function, which describes the delivery probability gained by prefetching.

The cost function is defined as:

$$f_{cost}[n] = f_{IAAP-St}[n] - f_{CONI-St}[n] \qquad (11)$$

where $f_{CONI-St}[n]$ is the storage ratio function (2) of CONI and $f_{IAAP-St}[n]$ is the storage ratio function of IAAP. Since the (2) is a discrete function, the x is converted to a expected encounter number by $\lceil x\lambda_M \rceil$ to produce a proximate value.

Moreover, let the gain function $f_{gain}(x)$ be:

$$f_{gain}(x) = -F_{CONI-D}(x) + F_{IAAP-D}(x) \qquad (12)$$

where $F_{CONI-D}(x)$ is the CDF of CONI-D [13] (9) and $F_{IAAP-D}(x)$ is the CDF of IAAP-D. The utility versus time with regards to different request rate R is shown in Fig.8b. The M is set as 7 while R^{-1} is set as 200s, 400s, 800s and 1600s accordingly. As the curves indicate, the maximum utility happens around 500s to 800s which suggests that the best value of t_r under above setup for (10). However, small t_r may result in low delivery probability. Since this is just an example, it would be better to take more factors into consideration.

VII. SIMULATION EXPERIMENTS

In this section, the validation of proposed models are carried out. To examine how well the proposed model can estimate the delivery process, various combination of the factors are tested and investigated against simulation results. The scalability and robustness are also testified.

The Riverbed Modeler [18] is adopted as simulation platform in this work. For every scenario, the same distribution of velocity and pause time are applied to every node, which means that each pair of nodes has the same pairwise encounter rate λ_p. The communication range r is set as 25m considering potential noises or obstacles between neighboring nodes, length of the side of area a is set as 500m. According to [13], although this setting isn't in line with the assumption that transmission radius r is much smaller than side length a, the encounters of three or more nodes are guaranteed to be in a small percentage between simulations and analytical models.
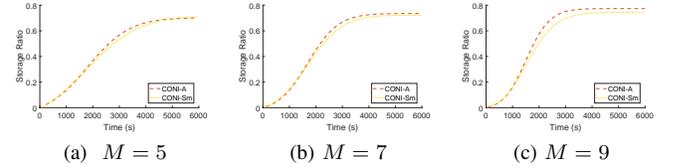
The content lifetime t_c is set as 12000s. The consumer starts requesting contents after 5000s to minimize the affect of initial node distribution.

The CONI and IAAP are implemented in Riverbed Modeler [18], and follow the description in Section III. The simulation provides a rather realistic data of the content delivery in opportunistic scenarios to verify the accuracy, scalability and robustness of proposed analytical models. Every simulation collects around 3000 sets of data against the corresponding performance factor.

A. Storage Ratio

The collection of simulated N_i starts when the request to content i is sent from the consumer to the first encountered node. The N_i is recorded every 60 seconds. The collection finishes when content i expires. The request lifetime t_r is set as 6000s. All the simulated data are averaged to produce the storage ratio.

First, the Fig.9 demonstrates the analytical CONI-St (CONI-A) compared to the simulated results of CONI (CONI-Sm). The data over 6000s are omitted because the requests expire after 6000s. To test the scalability of CONI-St, the total number of nodes M is set as 5, 7 and 9 and demonstrate in Fig.9 accordingly. The encounter rate of the whole network λ_M changes when M changes. The CONI-St shows good scalability when M varies. And the results clearly indicate that CONI-St estimates the simulation results well.

Then, the Fig.10 shows the comparison between analytical results (IAAP-A) produced by IAAP-St and simulated IAAP (IAAP-Sm). The CONI-A is also added as a reference. Under the assumptions stated in Section IV-D, the IAAP-St model performs the estimation of the number of prefetched contents ideally. In simulations which are closer to practical cases, the content i can be prefetched by request to content $i - 1$ after request to content i sent out. The reason causing this is that nodes cannot know when the next request is sent out, thus the prefetching cannot be disabled immediately after request to content i sent out.

To test the robustness of IAAP-St, the request interval R^{-1} is set as 400s, 800s, 1200s and 1600s and shown in Fig.10 respectively. The start point of the curve is the prefetched storage ratio. The IAAP-A shows good estimation of the prefetched storage ratio of IAAP-Sm. The simulated maximum storage ratio is smaller than the analytical one when R^{-1} is less or equal to 1200s. When R^{-1} larger than 1200s, the simulated maximum storage ratio is higher than the analytical one. The best estimation is achieved near $R^{-1} = 1200s$. Since both
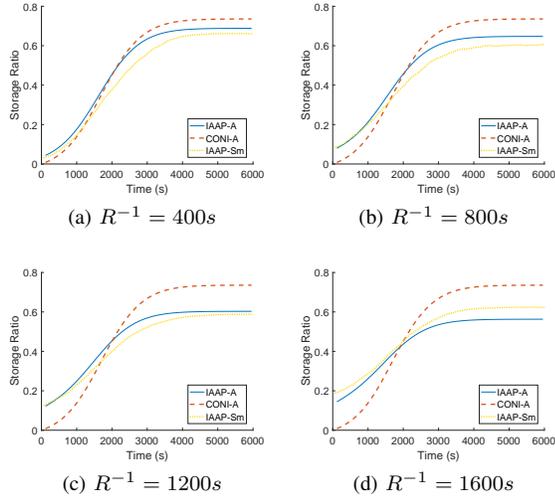
(a) $R^{-1} = 400s$

(b) $R^{-1} = 800s$

(c) $R^{-1} = 1200s$

(d) $R^{-1} = 1600s$

Fig. 10. Analytical IAAP-Storage ratio vs. analytical CONI-Storage ratio vs. simulation result: $t_r = 6000s$, $M = 7$

(a) $M = 5$ (b) $M = 7$ (c) $M = 9$

Fig. 11. Analytical CONI encounter number distribution vs. simulated CONI encounter number distribution: $R^{-1} = 1200s$, $t_r = 6000s$

(a) $M = 5$ (b) $M = 7$ (c) $M = 9$

Fig. 12. Analytical CONI-Delay distribution vs. simulated CONI-Delay distribution: $R^{-1} = 1200s$, $t_r = 6000s$

requests to content $i-1$ and request to content i will trigger the fetching of content i, the long R^{-1} allows requests cover different agents. Therefore, the differences between IAAP-A and IAAP-Sm are contents prefetched by requests to content $i-1$. Because the number of requests to content $i-1$ increases with R^{-1}, the differences in Fig.10d, which has largest R^{-1}, are larger compared to others. Meanwhile, when R^{-1} is smaller, the requests are likely to cover same agents, which reduces the content dissemination speed. In conclusion, the IAAP-St model can produce accurate prediction of prefetched storage ratio when $R^{-1} \leq 1200s$, and estimate the lower bound of IAAP storage ratio when $R^{-1} > 1200s$. Combined with the CONI-St estimating the upper bound of storage ratio after 3000 sec, IAAP-St can provide a comprehensive result of storage ratio with prefetching.

B. Delay Distribution

The end-to-end delay of contents are collected only when contents are delivered. Since the contents can only be requested by matched requests, once the requests are expired, contents will not be replicated any more. Thus, if the consumer fails to get contents within request lifetime t_r, the deliveries fail and are not recorded.

First, Fig.11 shows the comparison between the CONI-Sm and the CONI-A. Referring to the assumptions made in III, the t_r is set as 6000s to allow enough contents to be delivered. The R^{-1} is set as 1200s, which is larger than λ_M, to ensure the requests will not bind together. The M is set as 5, 7 and 9 to test the model's scalability. It is worth pointing out that the tail of the analytical delay distribution, formed by undelivered contents, is discarded, causing the whole distribution curve moves towards left. That is the reason why the peak of CONI-Sm of Fig.11c is on the left of the peak of CONI-A. The difference between analytical and simulation could be reduced by enlarge t_r. Meanwhile, most of the contents are delivered within a few encounters. It is clearly shown in Fig.11a that
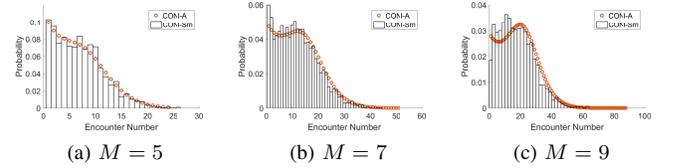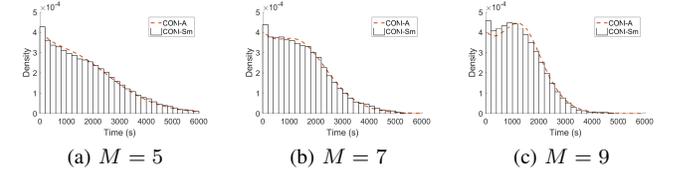
80% of deliveries happen within 10 encounters. Therefore, the left part of the distribution draws most of attentions.

For the delay distribution, data are collected in the same simulation as the encounter number test. The CONI-A shows little differences compared to the CONI-Sm as shown in Fig.12. All three cases shows good estimation of CONI-D [13]. Since the delay distribution can be interpreted as delivery probability vs. delay allowance, CONI-D [13] can help researchers find a certain delay distribution with finite delay rather than infinite delay.

When the second assumption made in Section III is relaxed, contents whose delay is longer than t_r can be excluded from distribution. The refined distribution is a truncated distribution with range $[0, t_r)$ which is shown below:

$$f_{trun}(x) = \frac{f_{mix}(x)}{F_{mix}(t_r) - F_{mix}(0)}$$

where the $F_{mix}(t_r)$ is the CDF of (9).

In Fig.13, to test whether CONI-D [13] can be adapted to truncated delay distribution, the t_r is set as 1200s, 2400s, 3600s, 4800s and 6000s and demonstrated respectively. The analytical Truncated CONI (Trun-CONI-A) is compared to the CONI-Sm, and the results indicate that truncated delay CONI-D [13] works well.

Finally, shown in Fig.14, the IAAP-A is compared to IAAP-Sm, and CONI-A is added as a reference. The R^{-1} is set as 400s, 800s, 1200s and 1600s to investigate the robustness of IAAP-D and shown in Fig.14 accordingly. Because the prefetching shortens the delay greatly, the CDF of delay distribution is adopted to make comparison easier than PDF. When R^{-1} increases, more contents are prefetched, the first column of IAAP-Sm increases dramatically. The accurate estimation of first column indicates that the estimation of initial distribution of IAAP is accurate. Although, the increasing rate of IAAP-A of Fig.14a and Fig.14b is a bit lower than IAAP-Sm, overall the IAAP-D estimates simulation well. According to the diagram, the differences between IAAP and CONI become smaller when time increases. Thus, the t_r should be shorten to reducer further cost.
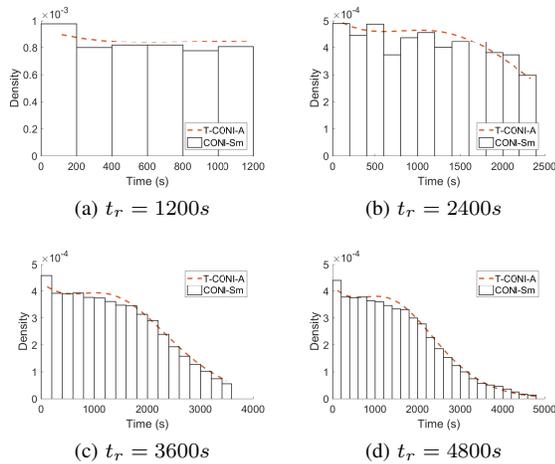
(a) $t_r = 1200s$ (b) $t_r = 2400s$

(c) $t_r = 3600s$ (d) $t_r = 4800s$

Fig. 13. Analytical truncated CONI-D distribution vs. simulated delay distribution, $R^{-1} = 100s$, $M = 5$

(a) $R^{-1} = 400s$ (b) $R^{-1} = 800s$

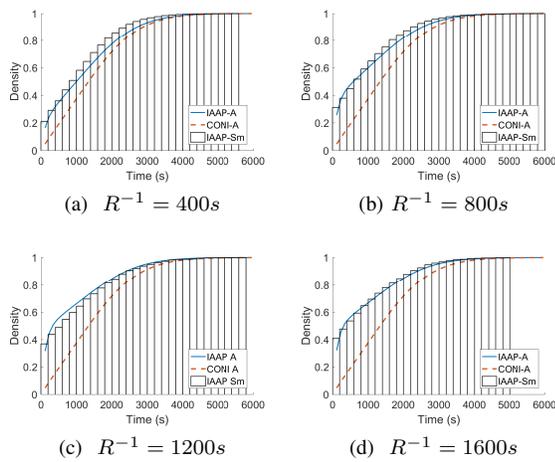(c) $R^{-1} = 1200s$ (d) $R^{-1} = 1600s$

Fig. 14. IAAP-Storage ratio vs. CONI-Storage ratio vs. simulation result: $t_r = 6000s$, $M = 7$

VIII. DISCUSSION AND FUTURE WORK

In this paper, several analytical models are proposed with certain assumptions, allowing correct understanding when applying the model in practical cases. In this section, the assumptions and corresponding restrictions are discussed below together with future works:

- Homogeneous mobility models. The homogeneous mobility models are applied to scenarios where there all the carrier of devices are treated as the same type, for instance, either disaster-relief team or vehicles, are treated equally. For scenarios consist of both humans and vehicles, the heterogeneous mobility models are preferred. Meanwhile, time variant mobility model are also discussed in some literatures. In this paper, the change of node number due to node failure is ignored to allow focusing on models. The heterogeneous mobility model and time variant mobility, together with node failure affect the models will be carried out in future work.

- Popular content prediction hit ratio. The prediction hit ratio is assumed to be 100% in this work to allow the

focus of the model analysis. However, the model can be extended to any hit ratio by applying the hit ratio to initial states distribution estimation. The extended initial states distribution estimation with deep learning based content prediction will be proposed in the future work.

IX. CONCLUSION

By investigating the IoT networks in opportunistic scenarios, the degradation of performance of communication affects the IoT's functions. In order to better understand the opportunistic scenarios, both the delay distribution and storage ratio of CONI in opportunistic scenarios are modelled. With four assumptions, the proposed CONI-D [13] and CONI-St fit the simulation results well. Moreover, the truncated CONI-D [13] could also expand to more realistic simulation cases. For simple cases, the proposed model can help accelerate the research process.

Although the mobility features are determined by the opportunistic scenarios, and are not tuneable from communication side, the improvement can be made by applying prefetching technologies. The models for IAAP is proposed to estimate the impact of prefetching over IoT in opportunistic scenarios. The IAAP-D can estimate the simulation results well while IAAP-St can predict the prefetched storage ratio. The IAAP consumes storages to reduce the delivery delay. For the future work, we would like to propose a prefetching scheme to balance the storage consumption and delay reduction dynamically.

REFERENCES

[1] Y. Shibata and G. Sato, "Iot based mobility information infrastructure in challenged network environment toward aging society," in *Advanced Information Networking and Applications Workshops (WAINA), 2017 31st International Conference on*. IEEE, 2017, pp. 645–648.

[2] R. Balakrishnan *et al.*, "Post disaster damage estimation using integrated gps sensor network & gis," in *Wireless Communications, Signal Processing and Networking (WiSPNET), International Conference on*. IEEE, 2016, pp. 1234–1238.

[3] Y. Cao and Z. Sun, "Routing in delay/disruption tolerant networks: A taxonomy, survey and challenges," *IEEE Communications surveys & tutorials*, vol. 15, no. 2, pp. 654–677, 2013.

[4] G. Tyson, J. Bigham, and E. Bodanese, "Towards an information-centric delay-tolerant network," in *Computer Communications Workshops (INFOCOM WKSHPS), 2013 IEEE Conference on*. IEEE, 2013, pp. 387–392.

[5] G. Xylomenos, C. N. Ververidis, V. A. Siris, N. Fotiou, C. Tsilopoulos, X. Vasilakos, K. V. Katsaros, and G. C. Polyzos, "A survey of information-centric networking research," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 2, pp. 1024–1049, 2014.

[6] D. Kutscher, S. Eum, K. Pentikousis, I. Psaras, D. Corujo, D. Saucez, T. Schmidt, and M. Waehlisch, "Information-centric networking (icn) research challenges," Tech. Rep., 2016.

[7] K. Pentikousis, B. Ohlman, D. Corujo, G. Boggia, G. Tyson, E. Davies, A. Molinaro, and S. Eum, "Information-centric networking: baseline scenarios," Tech. Rep., 2015.

[8] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking named content," in *Proceedings of the 5th international conference on Emerging networking experiments and technologies*. ACM, 2009, pp. 1–12.

[9] H. Jin, D. Xu, C. Zhao, and D. Liang, "Information-centric mobile caching network frameworks and caching optimization: a survey," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, p. 33, 2017.

[10] V. A. Siris, X. Vasilakos, and G. C. Polyzos, "Efficient proactive caching for supporting seamless mobility," *arXiv preprint arXiv:1404.4754*, 2014.

[11] R. Groenevelt, P. Nain, and G. Koole, "The message delay in mobile ad hoc networks," *Performance Evaluation*, vol. 62, no. 1, pp. 210–228, 2005.

[12] F. Neves dos Santos, B. Ertl, C. Barakat, T. Spyropoulos, and T. Turletti, "Cedo: Content-centric dissemination algorithm for delay-tolerant networks," in *Proceedings of the 16th ACM international conference on Modeling, analysis & simulation of wireless and mobile systems*. ACM, 2013, pp. 377–386.

[13] J. Yang, Y. Sun, and J. Requena Carrion, "Modelling the delay distribution of information centric delay tolerant networks in disaster area," in *MobileHealth 2017: 7th ACM MobiHoc Workshop on Pervasive Wireless Healthcare (MobileHealth'2017)*, Chennai, India, Jul. 2017.

[14] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Pocket switched networks and human mobility in conference environments," in *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*. ACM, 2005, pp. 244–251.

[15] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile computing*. Springer, 1996, pp. 153–181.

[16] S. Basagni, I. Chlamtac, V. R. Syrotiuk, and B. A. Woodward, "A distance routing effect algorithm for mobility (dream)," in *Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking*. ACM, 1998, pp. 76–84.

[17] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless communications and mobile computing*, vol. 2, no. 5, pp. 483–502, 2002.

[18] R. Technology, "Steelcentral riverbed modeler," 2018. [Online]. Available: https://www.riverbed.com/gb/products/steelcentral/steelcentral-riverbed-modeler.html

[19] A. Keränen, J. Ott, and T. Kärkkäinen, "The one simulator for dtn protocol evaluation," in *Proceedings of the 2nd international conference on simulation tools and techniques*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009, p. 55.

[20] T. Karagiannis, J.-Y. Le Boudec, and M. Vojnovic, "Power law and exponential decay of intercontact times between mobile devices," *IEEE Transactions on Mobile Computing*, vol. 9, no. 10, pp. 1377–1390, 2010.

[21] A. Vahdat, D. Becker *et al.*, "Epidemic routing for partially connected ad hoc networks," 2000.

[22] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*. ACM, 2005, pp. 252–259.

[23] R. Diana and E. Lochin, "Modelling the delay distribution of binary spray and wait routing protocol," in *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2012 IEEE International Symposium on a*. IEEE, 2012, pp. 1–6.

[24] A. Picu and T. Spyropoulos, "Dtn-meteo: Forecasting the performance of dtn protocols under heterogeneous mobility," *IEEE/ACM Transactions on Networking (TON)*, vol. 23, no. 2, pp. 587–602, 2015.

[25] C. Anastasiades and T. Braun, "Dynamic transmission modes to support opportunistic information-centric networks," in *Networked Systems (NetSys), 2015 International Conference and Workshops on*. IEEE, 2015, pp. 1–5.

[26] C. Anastasiades, T. Braun, and V. A. Siris, "Information-centric networking in mobile and opportunistic networks," in *Wireless Networking for Moving Objects*. Springer, 2014, pp. 14–30.

[27] T. Le, H. Kalantarian, and M. Gerla, "Socially-aware content retrieval using random walks in disruption tolerant networks," in *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2015 IEEE 16th International Symposium on a*. IEEE, 2015, pp. 1–6.

[28] L. Yao, A. Chen, J. Deng, J. Wang, and G. Wu, "A cooperative caching scheme based on mobility prediction in vehicular content centric networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 5435–5444, June 2018.

[29] A. Gellert and A. Florea, "Web prefetching through efficient prediction by partial matching," *World Wide Web*, vol. 19, no. 5, pp. 921–932, 2016.

[30] N. Abani, T. Braun, and M. Gerla, "Proactive caching with mobility prediction under uncertainty in information-centric networks," in *Proceedings of the 4th ACM Conference on Information-Centric Networking*. ACM, 2017, pp. 88–97.

[31] Y. Yao, Y. Sun, C. Phillips, and Y. Cao, "Movement-aware relay selection for delay-tolerant information dissemination in wildlife tracking and monitoring applications," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 3079–3090, 2018.

[32] A. Krifa, C. Barakat, and T. Spyropoulos, "Message drop and scheduling in dtns: Theory and practice," *IEEE Transactions on Mobile Computing*, vol. 11, no. 9, pp. 1470–1483, 2012.

Jinze Yang received a B.Eng. Degree in Internet of Things Engineering from Queen Mary University of London (QMUL) in 2015. He is currently a doctoral candidate in the school of Electronic Engineering and Computer Science, QMUL. His research interests include internet of things networks, information centric networks, ad hoc networks and mobility models.

Yan Sun received a B.Eng. Degree in Telecommunications Engineering from Beijing University of Posts and Telecommunications in 2001, and the MsC and Ph.D. degrees in Electronic Engineering from Queen Mary University of London in 2003 and 2009, respectively. She joined Siemens Ltd, China, in 2001 as a network optimization engineer and rejoined in 2003 as system engineer in R& D after obtaining her Master degree in UK. She then worked in Siemens (later Nokia Siemens Networks, Ltd.) as a product manager for 5 years, responsible for the 3rd generation Cellular Network equipment product lines. In 2009, Yan returned to Queen Mary as a lecturer. Her current research interests include Ad hoc networks, energy saving for modern mobile networks, software define networks and mobile healthcare networks.

Jess Requena-Carrin received the B.Sc. and M.Sc. degrees in telecommunications engineering and the Ph.D. degree from the Carlos III University of Madrid, Spain, in 2003 and 2008, respectively. From 2009 to 2014, he was a Lecturer with Rey Juan Carlos University, Madrid, Spain, and in 2014, he joined the School of Electronics Engineering and Computer Science, Queen Mary University of London. He is currently a Lecturer with the Queen Mary University of London, London, U.K., and the Deputy Director of the Joint Innovation Centre between the Queen Mary University of London and the Beijing University of Posts and Telecommunications, Beijing, China. His main research interests include statistical signal processing and computer modeling and simulation of biological systems and biomedical measurements.

Yue Cao received the PhD degree from the Institute for Communication Systems (ICS), University of Surrey, UK in 2013. He was the Research Fellow at ICS, University of Surrey, UK; Lecturer and Senior Lecturer at Department of Computer and Information Sciences, Northumbria University, UK; and has been the International Lecturer at School of Computing and Communications, Lancaster University, UK. His research interests focus on Intelligent Transport Systems. He is the Associate Editor of IEEE Access, KSII Transactions on Internet and Information Systems, IGI Global International Journal of Vehicular Telematics and Infotainment Systems.