An Improved Stochastic Modeling of Opportunistic Routing in Vehicular CPS

Deze Zeng, Member, IEEE, Song Guo, Senior Member, IEEE, Ahmed Barnawi, Member, IEEE, Shui Yu, Senior Member, IEEE, and Ivan Stojmenovic, Fellow, IEEE

Abstract—Vehicular Cyber-Physical System (VCPS) provides CPS services via exploring the sensing, computing and communication capabilities on vehicles. VCPS is deeply influenced by the performance of the underlying vehicular network with intermittent connections, which make existing routing solutions hardly to be applied directly. Epidemic routing, especially the one using random linear network coding, has been studied and proved as an efficient way in the consideration of delivery performance. Much pioneering work has tried to figure out how epidemic routing using network coding (ERNC) performs in VCPS, either by simulation or by analysis. However, none of them has been able to expose the potential of ERNC accurately. In this paper, we present a stochastic analytical framework to study the performance of ERNC in VCPS with intermittent connections. By novelty modeling ERNC in VCPS using a token-bucket model, our framework can provide a much more accurate results than any existing work on the unicast delivery performance analysis of ERNC in VCPS. The correctness of our analytical results has also been confirmed by our extensive simulations.

Index Terms—Vehicular Cyber-Physical System, Epidemic Routing, Random Linear Network Coding, Stochastic Analysis

1 INTRODUCTION

With the recent rapid development on information and communication technologies, information systems transit from the pure cyber space to a hybrid cyber-physical space. Cyber-Physical System (CPS) will transform how people interact with things around by allowing direct observation, coordination and manipulation of the physical world via cyber technologies [1]. Vehicular CPS (VCPS) integrates the sensing, computing and communication capabilities on vehicles to support a diversity of applications such as road ads, safety improvement, intelligent transportation, environment estimation, on-road infotainment, etc [2], [3]. Vehicle-to-vehicle (V2V) communication plays an important role in these VCPS applications. For example, in vehicular participatory sensing, the collected data shall be delivered to the “sink node”, e.g., road-side unit (RSU), for post-processing [4]–[6].

The major challenge in the development of VCPS is the intermittent network connectivity that a transmission happens only when two vehicles opportunistically come into the communication range of each other. To cater for such characteristic, the communication in VCPS is conducted in a “store-and-forward” manner [7], i.e., a packet is stored at one node and then forwarded to another using V2V communication when there is a transmission opportunity, without relying on an instant network connectivity. Such new scheme has been intensively studied and adopted by epidemic routing [7]–[9] as an efficient way to tackle the routing issue in VCPS.

The basic idea of epidemic routing is that a packet is greedily disseminated at each transmission opportunity in a hope that at least one copy will succeed in reaching its destination. Several variants (e.g., PRoPHET [10], MaxProp [8], RAPID [11], etc.) have been proposed with different goals, such as to maximize delivery rate, or minimize delivery delay and resources (e.g., buffer, energy) consumption. A major design challenging issue is that when disseminating multiple packets, a relay node can hardly make an optimal decision on which packet should be replicated and forwarded to achieve the best delivery performance in a network with unexpected connections. Fortunately, it has been found that network coding can tackle this issue and improve the performance significantly [12]–[15]. Using network coding, a node only needs to simply forward a random linear combination of packets it has received using random linear network coding technique upon each transmission opportunity. When the destination receives enough number of linearly independent coded packets, the original packets can be recovered. Epidemic routing using network coding (ERNC) can achieve much higher delivery performance than the non-coding schemes as shown by both simulations [12], [15] and analysis [13].

The existing theoretical work [13] is based on an assumption that each received coded packet at the destination is innovative with high probability. However, our simulation study shows that this assumption is not true.
and the resulting analysis only roughly approximates the reality. This motivates us to develop an improved stochastic model to analyze the delivery performance of ERNC in terms of delivery delay of a group of packets. In particular, we propose a token bucket model that can achieve a significantly improved accuracy for checking the independency of each received coded packet. To analyze such token-bucket based stochastic process, a two-dimensional time-heterogeneous Markov chain is developed, and then the analytical performance of ERNC is derived using ordinary differential equations (ODEs). The correctness and accuracy of our theoretical analysis is validated by extensive simulations. Thanks to the accurate capturing and modeling the network coding behaviours, our analysis shows much higher accuracy, compared to existing theoretical work.

The rest of this paper is organized as follows. Section 2 gives the system model and problem statement. Section 3 presents stochastic modeling and analysis on ERNC. Section 4 shows our performance evaluation results. Section 5 provides a brief overview of related work. Finally, Section 6 concludes our work. For the conveniences of the readers, the major notations used in this paper are listed in Table 1.

## 2 System Model and Problem Statement

### 2.1 Network Model

We consider a VCPS with $N + 1$ vehicles, i.e., mobile nodes. Two nodes are able to communicate only when they come into the reciprocal radio range and we define this as a “contact” between them. Let pairwise inter-contact interval between a pair of nodes denote the duration of time from the time when they go out of transmission range of each other to the next time they come into each other. To improve the existing analytical results on the performance of ERNC in VCPS, we adopt the same mobility model, in which the pairwise encounter interval satisfies the exponential distribution with the same rate $\lambda$. This model has been widely accepted in the literature [13], [16], [17] because it is validated as a good approximation for the inter-contact interval in a number of realistic vehicular networks [16], [18].

The packet size is the maximum data that can be transferred from one node to another during one contact. In other words, at most one packet is forwarded at each transmission opportunity. Similar to the models used in [13], [17], we assume a single-packet memory at each relay node and at most one packet is allowed to be carried by a relay node during its movement. We consider a unicast application session with $K$ packets originated from the source node and destined to the sole destination node via the help of relay nodes.

### 2.2 Routing Scheme

Since network partition happens in VCPS, all transmissions are performed in a “store-and-forward” manner. A packet is stored in the local buffer of its carrier until a transmission opportunity arises. ERNC scheme has been shown with much improved delivery performance over the traditional non-coding ones by both empirical study [15] and theoretical analysis [13]. In ERNC, a coded packet $p^{\text{coded}}$ is a linear combination of native packets $p_1^{\text{native}}, p_2^{\text{native}}, \ldots, p_K^{\text{native}}$ in the form: $p^{\text{coded}} = \sum_{i=1}^{K} \alpha_i p_i^{\text{native}}$, where $\alpha_i, i = 1, \ldots, K$, are called coding coefficients and are randomly chosen from a Galois Fielded ($\mathbb{F}_q$) [12]. The source node generates such a coded packet by randomly encoding all the native packets and forwards it to the encountered node, which shall reserve the coded packet in its local buffer. Suppose two relay nodes $a$ and $b$ encounter and a coded packet $p_a^{\text{coded}}$ is forwarded from node $a$ to node $b$. Once received, node $b$ updates its own encoded packet $p_b^{\text{coded}}$ as $p_b^{\text{coded}} = \exp_b^{\text{coded}} + \beta b^{\text{coded}}$, where $\alpha$ and $\beta$ are also randomly chosen from $\mathbb{F}_q$. In this way, single buffer that holds one coded packet at each relay node can achieve almost the same performance as the multiple buffer case as discovered in [13]. After the destination node collects an enough number of linearly independent coded packets, it is able to retrieve all the original native packets by Gaussian Elimination.

Let us use the example shown in Fig. 1 to illustrate the ERNC process, in which source node $s$ has two packets $p_1$ and $p_2$ to be delivered to destination node $d$.

#### $t_1$

Source $s$ encounters relay $r_1$ and transmits a linearly coded packet $x = 1p_1 + 2p_2$ with coefficients randomly selected from $\mathbb{F}_{256}$. Suppose the buffer of $r_1$ is empty, its received coded packet is stored in its original form.

#### $t_2$

Relay $r_1$ encounters destination $d$ and forwards $x$ in its buffer.

### Table 1: Notations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>number of nodes in the network</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>pair-wise contact rate</td>
</tr>
<tr>
<td>$K$</td>
<td>the total number of packets to deliver</td>
</tr>
<tr>
<td>$\mathbb{F}_q$</td>
<td>Galois field with field size $q$</td>
</tr>
<tr>
<td>$P_{\text{inf}}(t)$</td>
<td>The number of nodes that are infected with coded packets</td>
</tr>
<tr>
<td>$p^{\text{inf}}(k,t)$</td>
<td>The probability that the destination node receives $k$ coded packets</td>
</tr>
<tr>
<td>$P_{\text{all}}(K,t)$</td>
<td>The probability that the destination node receives at least $K$ coded packets</td>
</tr>
<tr>
<td>$I_{\text{inf}}(t)$</td>
<td>The number of nodes that are infected with one token</td>
</tr>
<tr>
<td>$p_{\text{in}}(t)$</td>
<td>The probability that the destination node receives one token</td>
</tr>
<tr>
<td>$P_{\text{out}}(K,t)$</td>
<td>The probability that the destination node receives at least $K$ tokens</td>
</tr>
<tr>
<td>$\lambda(t)$</td>
<td>The packet arrival rate in the token bucket model</td>
</tr>
<tr>
<td>$\mu(t)$</td>
<td>The token arrival rate in the token bucket model</td>
</tr>
<tr>
<td>$X(N_s,N_v,t)$</td>
<td>The probability that the destination node receives $N_s$ packets and $N_v$ tokens</td>
</tr>
<tr>
<td>$E[T]$</td>
<td>The expected delivery delay of $K$ packets using ERNC</td>
</tr>
</tbody>
</table>
In other words, a significant portion of received coded packets are actually dependent, i.e., non-innovative. This further incurs inaccuracy of the analysis on delivery delay. In Fig. 2(b), we show both simulation and analytical results (according to [13]) of delivery delay in cumulative distribution function (CDF). We observe that the analytical result is quite optimistic compared to the simulated one. For example, the decoding is expected at time $110$ with probability $70\%$ by analysis, while the true probability is only around $30\%$. Therefore, the key challenging factor to the performance evaluation of ERNC in VCPS is that an accurate stochastic modeling of both data transmission and decoding is required. This motivates us to find an accurate analytical model that can truly reveal the successful decoding probability at a given time.

3 Stochastic Modeling and Analysis on ERNC

In this section, we first present a logical view of the decoding process and describe it by a token-bucket model. We then develop a two-dimensional Markov chain based on the token-bucket model to derive the probability of decoding $K$ native packets at the destination node.

3.1 Decoding Process Modeling

By ERNC, each coded packet in $p^{\text{coded}} = [p_1^{\text{coded}}, \ldots, p_K^{\text{coded}}]$ is a linear combination of $K$ native packets $p^{\text{native}} = [p_1^{\text{native}}, \ldots, p_K^{\text{native}}]$. The destination node will successfully decode all native packets when it is able to solve a linear equation system $p^{\text{coded}} = C p^{\text{native}}$, where $p^{\text{native}}$ are unknown variables and $C$ is the coefficient matrix generated by the ERNC process. As long as $K$ innovative coded packets are received, all $K$ native packets are decodable. Conventionally, to determine the innovativeness of a coded packet is done by checking whether the received coded packet can increase the rank of all received ones so far. However, it is difficult to directly model the coefficient matrix because $p^{\text{native}}$ is randomly encoded into $p^{\text{coded}}$ by relay nodes during contacts.

We observe that at any time, the rank of all coded packets in the network is equal to the number of native packets that have been injected into the network by the source. The rank of all received coded packets by the destination can thus be obtained by checking the number of “received” unknown variables that have shown in the corresponding linear equation system. Specifically, the reception of an unknown variable implies that the corresponding coefficient is non-zero in at least one equation received so far at the destination. Therefore we define such arrival of “unknown variable” as virtual arrival to distinguish from the physical arrival of coded packets. This leads to a simple way to check the innovativeness of a coded packet as follows:

Lemma 1: If a coded packet makes the number of equations exceed the number of unknown variables at
Lemma 1 is a necessary but not sufficient condition for decoding. Our simulation study shows that a process only applying the rule in Lemma 1 can approximate the realistic decoding process with a very high accuracy.

To validate the proposed method, we conduct simulation experiments with the following rules for innovativeness checking.

- **Conventional rule**: Upon receiving a coded packet from an encountered node, the destination node checks if its coefficient vector is linearly independent to the coefficient matrix of other packets that have been already received (i.e., rank checking). If and only if it is innovative, it is accepted. Otherwise, it is discarded.

- **Optimistic rule**: According to [13], it is assumed that the received coded packet is always innovative. As a result, without any independency checking, any coded packet is always accepted unconditionally.

- **Lemma-1 based rule**: A coded packet is accepted if and only if the total number of received unknown variables exceeds the number of coded packets received so far.

Compared with the conventional rule, if both rules regard a coded packet as non-innovative, we say our rule hits while if a non-innovative coded packet is treated as innovative by our rule, it is a miss. The simulation of delivering 100 packets is conducted in a network with $N = 200$ and $\lambda = 0.005$. The statistics of hit ratio and miss ratio are obtained by running it 5000 times. We notice that our rule can achieve 100% hit ratio and very low miss ratio, which is plotted in CDF as shown in Fig. 3. For example, most cases by Lemma 1 have a miss ratio of 0% and 97.7% of them are with a miss ratio less than 5% while only 65.4% of optimistic model achieves such miss ratio. This validates the high accuracy of our innovativeness checking rule.

Our experimental results also show that the processes of physical arrival and virtual arrival can be approximately considered as independent since their Pearson product-moment correlation coefficient [19] and mutual information [20] are as low as -0.039 and 0.0154, respectively, by running simulations 10000 times using different random seeds. These discoveries further inspire us to describe the decoding process by a token bucket model with two independent arrival processes as shown in Fig. 4. It works as follows.

1) A token arrival is equivalent to a virtual reception. Tokens are collected in an unlimited buffer.

2) A physical (packet) arrival can be accepted, called
3.2 Analysis on the Physical Arrival Process

According to the process of ERNC, a packet arrival is equivalent to a contact between the destination node and a node with coded packet. Under the infectious disease model, if a node carries a packet in its local buffer, it is regarded as “infected”. Initially, only the source node is infected and all the other nodes are “susceptible” to be infected. Once the transmission starts, the source node begins to infect the other nodes by the contact opportunities during its movement. The infected nodes (excluding the destination) are then able to infect other susceptible nodes and so on. Therefore, the number of infected nodes at time \( t \) can be described by the following ordinary differential equation (ODE):

\[
\frac{dI^{phy}(t)}{dt} = \lambda I^{phy}(t)(N - I^{phy}(t)), \quad (1)
\]

where \( \lambda \) is the average pairwise contact rate.

Solving (1) under the initial condition \( I^{phy}(0) = 1 \), (i.e., at time 0, only the source node can be regarded as infected), we get

\[
I^{phy}(t) = \frac{N}{1 + (N - 1)e^{-\lambda N t}}. \quad (2)
\]

Next, we can derive the expected number of infected nodes that have contacted the destination node by time \( t \) as:

\[
g(t) = \int_0^t \lambda I^{phy}(\tau)d\tau = \ln e^{\lambda N t} + N - 1. \quad (3)
\]

In our VCPS model, the contact process between any two nodes is a Poisson process since their contact interval is exponentially distributed. The contact process of meeting a infected node by the destination node is the sum of the Poisson processes with each destination node and is also a Poisson process with rate equals to these Poisson processes’ rate. In other words, the destination node contacts an infected node in an exponentially distributed interval. As a result, the probability that there are exactly \( k \) (\( k = 0, 1, 2, \cdots \)) contacts with the infected nodes by time \( t \) can be written as

\[
p^{phy}(k,t) = \frac{g(t)^k}{k!}e^{-g(t)}. \quad (4)
\]

Therefore, the probability that there are at least \( K \) packet arrivals at the destination node at time \( t \) is:

\[
p^{phy}(K,t) = 1 - \sum_{k=0}^{K-1} \frac{g(t)^k}{k!}e^{-g(t)} = 1 - \frac{N}{e^{\lambda N t} + N - 1} \sum_{k=0}^{K-1} \frac{(\ln e^{\lambda N t} + N - 1)^k}{k!}. \quad (5)
\]

It shows that the packet arrival process is a non-homogeneous Poisson process.

3.3 Analysis on the Virtual Arrival Process

A token, say \( t_i \), emerges in the network whenever the source node delivers a native packet \( (p_i^{native}) \) to another node encountered. In ERNC, any relay node is “infected” by this token as long as the coded packet in its buffer has a non-zero coefficient of the component relating to \( p_i^{native} \). The destination node detects the arrival of token \( t_i \) when it, by the first time, encounters a relay node infected by \( t_i \). Let \( T^{vir}(t) \) be the number of relay nodes that have been infected by token \( t_i \) at time \( t \) with a starting point \( t = 0 \) when the source node delivers \( p_i^{native} \). A fundamental difference from the last section is that after \( t = 0 \), the source node is get ready for the delivery of next token and only the encountered node that receives \( t_i \) can be regarded as infected, i.e., \( T^{vir}(0) = 1 \). Therefore, we have

\[
\frac{dT^{vir}(t)}{dt} = \lambda I^{vir}(t)(N - 1 - T^{vir}(t)). \quad (6)
\]

Solving (6) under the initial condition \( T^{vir}(0) = 1 \) yields

\[
T^{vir}(t) = \frac{N - 1}{1 + (N - 2)e^{-\lambda N t} - 1}. \quad (7)
\]

Let \( T^{vir}_p \) be the propagation delay of token \( t_i \), which is defined as the time from the source delivers \( p_i^{native} \) to the time when the destination detects its arrival, i.e., encounters a relay node infected by \( t_i \) by the first time. Since the propagation of \( t_i \) is epidemic, the CDF of \( T^{vir}_p \), denoted as \( P^{p}_p(t) = \Pr(T^{vir}_p < t) \), can thus be obtained by solving the following ODE developed in [21]:

\[
\frac{dP^{p}_p(t)}{dt} = \lambda I^{vir}(t)(1 - P^{p}_p(t)). \quad (8)
\]
Notice that in our model, packet $p_{\text{native}}$ (equivalently token $t_k$) could be directly delivered to the destination at the very beginning with probability $1/N$, i.e., $P_{p}^{t_{\text{vir}}} (0) = 1/N$.

By solving the ODE in (7), we have:

$$ P_{p}^{t_{\text{vir}}} (t) = 1 - C_1 e^{- \int \lambda t_{\text{vir}} (t) dt}, $$

where $\int \lambda t_{\text{vir}} (t) dt$ can be derived as follows:

$$ \int \lambda t_{\text{vir}} (t) dt = \int \frac{\lambda (N - 1)}{1 + (N - 2) e^{\lambda (N - 1) t}} dt = \int \frac{\lambda e^{\lambda (N - 1) t}}{N - 2 + e^{\lambda (N - 1) t}} dt = \ln(N - 2 + e^{\lambda (N - 1) t}) + C_2. $$

Let $C_3 = C_1 e^{-C_2}$, $P_{p}^{t_{\text{vir}}} (t)$ can be written as:

$$ P_{p}^{t_{\text{vir}}} (t) = 1 - \frac{1}{N - 2 + e^{\lambda (N - 1) t}}. $$

Solving $P_{p}^{t_{\text{vir}}} (0) = 1/N$ yields:

$$ C_3 = \frac{(N - 1)^2}{N}. $$

Taking (11) into (10), we finally obtain the solution of $P_{p}^{t_{\text{vir}}} (t)$ from (7):

$$ P_{p}^{t_{\text{vir}}} (t) = 1 - \frac{(N - 1)^2}{N(N - 2 + e^{\lambda (N - 1) t}).} $$

Now we consider the system starting time $t = 0$ that begins when the source node is just ready for data delivery. Each token goes into network with an exponentially distributed interval at rate $\lambda N$ according to the routing protocol. Thus the probability that the $k$-th token arrives in the network at time $t$ is equivalent to the probability that total $k$ tokens have been injected, as a Poisson process, into the network at time $t$:

$$ p_a (k, t) = \frac{(\lambda N t)^k}{k!} e^{-\lambda N t}. $$

Then the $k$-th token propagates to the destination node independently in an epidemic manner. The CDF of its arrival time, denoted as $P_{p}^{t_{\text{vir}}} (k, t)$, can be calculated as

$$ P_{p}^{t_{\text{vir}}} (k, t) = \int_0^t p_a (k, \tau) P_{p}^{t_{\text{vir}}} (t - \tau) d \tau. $$

Simulation based study is also conducted to verify the correctness of (14). In our experiments, we first specify any token, and then trace and check the time when a coded packet including such token is first received by the destination. We run the simulation several rounds with different random seeds to obtain the CDFs of arrival time for the 50th token and 100th token. Both results from simulation and (14) are shown in Fig. 5. They match quite well in both cases. This confirms our understanding that each token can be regarded as propagating independently and individually by epidemic routing, as if no other tokens exist.

We notice that the arrival of the $k$-th token at the destination node does not mean the arrival of at least $k$ tokens because their propagations are independent. According to our token bucket model, we are interested in the probability that at least $k$ tokens arrive at the destination node by time $t$. Recall that the emergence of tokens is a Poisson process. By the uniformity property of Poisson process, the time that a token appears in a given period is uniformly distributed.

Supposing the time that such token appears in the network is $\tau$ and its arrival time at the destination is $t$ (i.e., the propagation delay is $t-\tau$), we apply the uniform property to derive the probability that each token has arrived at the destination node at time $t$ as:

$$ p_a (t) = \left( \int_0^t \frac{(N - 1)^2}{N(e^{\lambda (N - 1)(t-\tau)} + N - 2)} d \tau \right) / t = 1 - \left( \int_0^t \frac{(N - 1)^2}{N(e^{\lambda (N - 1)(t-\tau)} + N - 2)} d \tau \right) / t = 1 - \frac{(N - 2) \lambda N t}{N - 1} \ln \left( \frac{e^{\lambda (N - 1) t} + e^{\lambda (N - 1) t} / (N - 2)}{N - 1} \right) = 1 - \frac{(N - 2) \lambda N t}{(N - 2) e^{-\lambda N t} + 1}. $$

Lemma 2: The probability that there are exactly $k$ token arrivals by time $t$ is

$$ p_d (k, t) = \frac{(\lambda N t)^k}{k!} e^{-\lambda N t}. $$

Proof: If $M (M \geq k)$ tokens have emerged during interval $(0, t)$, the probability that $k$ of them have arrived at the destination node by time $t$ is

$$ \binom{M}{k} p_a (t)^k (1 - p_a (t))^{M - k}, 0 \leq k \leq M, $$

where $p_a (t)$ is the probability that a token received by the destination node at time $t$, as defined in (17).
Note that the destination node can obtain \( k \) packets only when there are at least \( k \) tokens emerging in the network. In other words, \( M \) must be larger or equal than \( k \) (i.e., \( M = k, k+1, k+2, \ldots \)). As we have known, the tokens are sent out by the source node to the network following a Poisson process. The probability that there are \( M \) tokens in the network is \( p_M(M, t) \) as defined in (13). Therefore, we have:

\[
p_d^{vir}(k, t) = \sum_{M=k}^{\infty} \frac{(\lambda N t)^M}{M!} e^{-\lambda N t} \left( \frac{M}{k} \right) p_a(t)^k (1 - p_a(t))^{M-k} \]

(18)

Let \( Y = M - k, p_d^{vir}(k, t) \) can be then rewritten as:

\[
p_d^{vir}(k, t) = \frac{(\lambda N t p_a(t))^k}{k!} e^{-\lambda N t} \sum_{Y=0}^{\infty} \frac{(\lambda N t(1 - p_a(t)))^Y}{Y!} \]

(19)

From Lemma 2, we can see that the token arrival process is a heterogeneous Poisson process. This immediately leads to the following lemma.

**Lemma 3:** The probability that there are at least \( K \) token arrivals is:

\[
P_d^{vir}(K, t) = 1 - \sum_{i=0}^{K-1} p_d^{vir}(i, t). \]  

(20)

### 3.4 Analysis on the Decoding Probability

According to our analysis, both token and packet arrival processes are heterogeneous Poisson process with arrival rate

\[
\lambda(t) = \frac{d(\lambda N t p_a(t))}{dt} = \lambda N - \frac{\lambda + \lambda N(N-2)}{N - 2 + e^{-\lambda(N-1)t}} 
\]  

(21)

and

\[
\mu(t) = \frac{\lambda N}{1 + (N-1)e^{-\lambda N t}}. \]  

(22)

respectively. The state transition process can be described by a two-dimensional Markov chain shown in Fig. 6, where each state is represented by \( (N_p, N_v) \) and \( N_p \leq N_v \) holds. By denoting \( X(N_p, N_v, t) \) as the probability that the Markov chain stays at \( (N_p, N_v) \) at time \( t \), we can describe the transition process by a group of ODEs as follows. According to the transition characteristics, the whole figure can be divided into four regions.

**Region I:** all states in the first line except \((0, K)\). In Region-I, besides self-transition, state \((0,0)\) transits only to \((0,1)\) while all others except \((0,0)\) transits to two neighboring states. These can be described as:

\[
X'(0,0, t) = -\lambda(t)X(0,0, t) \]

\[
X'(0, i, t) = \lambda(t)X(0, i-1, t) - (\lambda(t) + \mu(t))X(0, i, t), \quad \forall i = [1, 2, \ldots, K-1] \]

(23)

**Region II:** all states in the diagonal line except \((0,0)\) and \((K, K)\). No packet arrival can be accepted any more if the system is in a state of Region-II as this is equivalent to the condition that all tokens in the token buffer have been used up already. Therefore, each state \((i, j)\) in Region-II depends on state \((i-1, j)\) and would transit to \((i, j+1)\). We then have

\[
X'(i,i, t) = \mu(t)X(i-1,i, t) - \lambda(t)X(i,i, t), \quad \forall i = [1, 2, \ldots, K-1] \]

(24)

**Region III:** all states in the last column. In a similar way, we can derive ODEs for the states in Region-III as:

\[
X'(0, K, t) = \lambda(t)X(0, K-1, t) - \mu(t)X(0, K, t), \]

\[
X'(i, K, t) = \lambda(t)X(i, K-1, t) + \mu(t)X(i+1, K, t), \quad \forall i = [0, 1, \ldots, K-1], \]

\[
X'(K, K, t) = \mu(t)X(K-1, K, t). \]

(25)

**Region IV:** all remaining states in the central region. Each state \((i, j)\) in Region-IV relies on its two neighboring states \((i-1, j)\) and \((i, j-1)\) on the upper and left, and
Fig. 7. Performance evaluation on the accuracy of the analysis

also would transit to the other two neighboring states 
(i + 1, j) and (i, j + 1) on the lower one and right:

$$X'(i, j, t) = \lambda(t)X(i, j - 1, t) + \mu(t)X(i - 1, j, t)$$
$$- (\lambda(t) + \mu(t))X(i, j, t),$$
$$\forall i = 1, 2, \ldots, K - 1, j = i + 1, \ldots, K - 1.$$  

(26)

Numerical method can be applied to solve the ODEs in (23)-(26) under the boundary condition:

$$X(i, j, 0) = \begin{cases} 
1, & (i, j) = (0, 0), \\
0, & \text{otherwise}. 
\end{cases}$$  

(27)

Note that, if the system transits into state (K, K), we conclude that all K native packets are decodable now. Therefore, X(K, K, t) represents the probability that all K packets have been successfully received by time t. In other words, the value of X(K, K, t) can be regarded as the CDF of decoding probability for K packets at time t.

Theorem 1: The expected delivery delay of K packets by ERNC in VCPS is

$$E[T] = \int_0^\infty (1 - X(K, K, \tau))d\tau,$$  

(28)

where X(K, K, \tau) is obtained by jointly solving (21)-(27).

4 PERFORMANCE EVALUATION

In this section, we present our evaluation on the accuracy of the delivery delay analysis by comparing the simulation results to the analysis results obtained by numerical method. The simulation results reported in this paper are obtained through a self-developed discrete-event simulator that strictly follows the system model presented in Section 2. For practical network coding operations, Galois field F_{256} is adopted. Numerical results are acquired using Matlab ODE solver ode45.

4.1 On the Accuracy of Our Analysis

According to our analysis, it can be inferred that the delivery probability of a content in VCPS is determined by its content size K (i.e., the number of packets), network size N and contact rate \lambda. In order to thoroughly show the accuracy of our analysis, we conduct three groups of simulations with different values of K, N and \lambda, respectively. In each group, in order to produce smooth CDF curve on the successful delivery probability on time t, 1000 rounds of simulations with different random seeds are performed. Fig. 7 shows the evaluation results in terms of CDFs of delivery delays and average delivery delay.

The parameter settings for our experiments are as follows: N = 200 and \lambda = 0.005 in Fig. 7(a), K = 100 and \lambda = 0.005 in Fig. 7(b), and N = 100 and K = 50 in Fig. 7(c). For all cases, the analytical and simulation results
are almost overlapped. The correctness high accuracy of our analysis is validated. This validates the correctness and accuracy of our stochastic modeling and analysis including the token bucket model for decoding process and the physical/virtual arrival process characterized by a time-dependent two-dimensional Markov chain.

To validate the accuracy of our analysis on average delivery delay as shown in Theorem 1, simulation studies under various network settings by differencing the values of $N$ and $\lambda$ (i.e., $(N, \lambda)=(100, 0.005), (100, 0.008)$ and $(200, 0.005)$) are conducted. Fig. 7(d) shows the average deliver delay $E[T]$ as a function of the number of packets ($K$) to deliver under each network setting. It can be obviously noticed that $E[T]$ is a linearly increasing function of $K$. Furthermore, we also see that the analysis results always tightly match with the simulation results. This validates the high accuracy of our derivation in Theorem 1.

4.2 On the Significance of Our Analysis

To show the significance of our accurate analysis, we consider a scenario where the content to be delivered is with a predetermined lifetime $T_i$; set according to the desired delivery probability, i.e., $T_i = \arg \min_l \Pr(b_l) \geq P$, where $\Pr(b_l)$ is the probability that the content has been decoded by time $t$ and $P$ is the desired delivery probability within the content’s lifetime. An outage happens when a content cannot be successfully decoded at the destination within its lifetime. We simulate the unicast of a content with 100 packets (i.e., $K=100$) in a network with $N=200$ and $\lambda = 0.005$. We vary the desired delivery probability and set the content lifetime according to analysis in [13] and ours (i.e., $\Pr(b_l) = X(K, K, t)$), respectively. For each case, 1000 simulation instances are conducted to obtain the outage probability. Fig. 8 shows the simulation results. Obviously, we can see that the inaccurate analysis on the decoding probability leads to a high outage probability. For example, when the desired delivery probability is 0.80, the outage probability is as high as 0.613 if the message lifetime is set according to [13]. Thanks to the accurate analysis on the decoding probability, our analysis can always make sure that the desired delivery probability is achieved.

5 RELATED WORK

5.1 Vehicular Cyber-Physical System

VCPS has emerged as a cutting edge technology for the next-generation intelligent transportation system. In the literature, various VCPS architectures have been proposed. Kim et al. [22] propose a CPS application framework that is able to provide a generic service to represent, manipulate, and share knowledge across the network without persistent network connectivity, e.g., Delay-Tolerant networks (DTNs). Jia et al. [3] propose a platoon-based VCPS architecture and investigate the expected time of safety message delivery among platoons with the consideration of traffic flow, vehicle, platoon size and transmission range. Ni et al. [23] explore a distributed cyber-physical solution by integrating networked and embedded sensing, computational intelligence and real-time vehicular communications to increase the safety and efficiency of transportation systems.

Besides the architecture issue, a diversity of algorithms have been studied towards different VCPS application optimizations. Miloslavov and Veeraraghavan [2] propose a sensor data fusion algorithm for calculating average per-division (i.e., a small segment of a route) speed and average route travel time in VCPS. Li et al. [24] devise a human factor aware service delivery scheme for VCPS that can transmit multiple messages with time-dependent utility to a subset of intended drivers so as to minimize the network-wide service utility loss. Later, Li et al. [25] study the targeted on-road ad delivery problem with the joint consideration of message scheduling and AP bandwidth allocation.

5.2 Content Delivery in DTNs

Since the underlying network of VCPS is of intermittent connectivity, the communications must be robust against delays and disruptions due to vehicle mobility. Much effort has been devoted to the performance analysis of DTN routing protocols.

For the delivery of a single packet, flooding-based epidemic routing, firstly proposed by Vahdat et al. [7], achieves the shortest delay [26]. Zhang et al. [21] derive the closed-form expression of delivery delay for single-packet unicast using epidemic routing in DTNs by an ODE-based approach. Subramanian et al. [27] analyze the efficiency of multi-hop routing over two-hop routing for multiple unicasts in DTNs with finite-sized buffer. Gunasekaran et al. [28, 29] apply Queueing Petri-Net and derive the end-to-end delivery delay from the modeled semi-Markov process. Khouzani et al. [30] investigate an energy-aware optimal epidemic routing in DTNs by a threshold-based scheme, where the forwarding probability is determined by relay node’s current remaining energy. Recently, Kim et al. [31] present a distributed cross-layer monitoring and optimization
method for secure content delivery toward decentralized content-based mobile ad hoc networking. Abdrabou et al. [6] analyze the message delivery delay in a two-lane road where vehicles moving in one direction send their sensing data to vehicles in the opposite direction in order to deliver the packets to the nearest road-side unit.

For the multi-packet delivery, network coding has been proved as a promising way to improve the delivery performance. There are many different network coding techniques available in the literature. Rateless codes, like LT code and Raptor code, require a predefined coding structure (e.g., Soliton distribution) and hence are not suitable for multihop forwarding. As a result, the studies on rateless codes, such as [17], [32]–[34], usually limit to two-hop routing (i.e., source-relay-destination), without the relay-to-relay transmissions.

When relay-to-relay transmission is considered, random linear network coding has shown its compelling strength in the delivery performance. Widmer et al. [12] compare the performance of random linear network coding based forwarding to alternative protocols and find out that it significantly outperforms probabilistic routing, particularly in challenging scenarios where connectivity is rare. In [15], Zhang et al. show that the network coding scheme achieves slightly smaller average block delay than non-coded schemes under unconstrained buffer case, but significant benefits under constrained buffer case. To optimize the network coding performance in DTNs, Ali et al. [35] propose a content delivery mechanism by combining network coding and global selective acknowledgement. The applicability of network coding to DTNs has been also validated by practical testbed. In [36], Joy et al. develop a network architecture using Android phones that exploits partial caches by utilizing network coding to deliver large files (e.g., images).

ACKNOWLEDGEMENT

This research is supported by the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) (Grant No. CUG140615). This research of Ivan Stojmenovic is partially supported by NSERC Discovery grant (Canada), and TR32016, Serbian Ministry of Science and Education (associated with Singidunum University, Serbia).

REFERENCES


Deze Zeng received his Ph.D. and M.S. degrees in computer science from University of Aizu, Aizu-Wakamatsu, Japan, in 2013 and 2009, respectively. He received his B.S. degree from School of Computer Science and Technology, Huazhong University of Science and Technology, China in 2007. He is currently an associate professor in School of Computer Science, China University of Geosciences (Wuhan), China. His current research interests include: cloud computing, software-defined sensor networks, data center networking, networking protocol design and analysis. He is a member of IEEE.

Song Guo (M102-SM11) received the Ph.D. degree in computer science from University of Ottawa, Canada. He is currently a Full Professor at School of Computer Science and Engineering, the University of Aizu, Japan. His research interests are mainly in the areas of protocol design and performance analysis for computer and telecommunications networks. He received the Best Paper Awards at ACM IMCOM 2014, IEEE CSE 2011, and IEEE HPCC 2008. Dr. Guo currently serves as Associate Editor of IEEE Transactions on Parallel and Distributed Systems and IEEE Transactions on Emerging Topics in Computing. He is in the editorial boards of ACM/Springer Wireless Networks, Wireless Communications and Mobile Computing, and many others. He has also been in organizing and technical committees of numerous international conferences, including serving as a General Co-Chair of MobiQuitous 2013. Dr. Guo is a senior member of the IEEE and the ACM.

Ahmed Barnawi is an associate professor at the Faculty of Computing and IT, King Abdullah University, Jeddah, Saudi Arabia, where he works since 2007. He received Ph.D at the University of Bradford, UK in 2006. He was visiting professor at the University of Calgary in 2009. His research areas are cellular and mobile communications, mobile ad hoc and sensor networks, cognitive radio networks and security. He received three strategic research grants and registered two patents in the US. He is IEEE member.

Shui Yu received the B.Eng. and M.Eng. degrees from University of Electronic Science and Technology of China, Chengdu, P. R. China, in 1993 and 1999, respectively, and the Ph.D. degree from Deakin University, Victoria, Australia, in 2004. He is currently a Senior Lecturer with the School of Information Technology, Deakin University, Melbourne, Australia. He has published around 100 peer review papers; many of them are in top journals and top conferences, such as IEEE TPDS, IEEE TIFS, IEEE TFS, IEEE TMC, and IEEE INFOCOM. His research interests include networking theory, network security, and mathematical modeling. Dr. Yu actively serves his research communities in various roles, which include the editorial boards of IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Mobile Computing, Elsevier Computer Networks, and three other international journals, TPC of IEEE INFOCOM, symposium co-chairs of IEEE ICC 2014, IEEE ICNC 2013 and 2104, and many different roles of international conference organizing committees. He is a senior member of IEEE, and a member of AAAS.
Ivan Stojmenovic was editor-in-chief of IEEE Transactions on Parallel and Distributed Systems (2010-3), is Associate Editor-in-Chief of Tsinghua Journal of Science and Technology, and is founder of four journals (Journal of Multiple Valued Logic and Soft Computing, Ad Hoc & Sensor Wireless Networks, International Journal of Parallel, Emergent and Distributed Systems, Cyber Physical Systems). He is editor of IEEE Transactions on Computers, IEEE Network, IEEE Transactions on Cloud Computing etc. Stojmenovic has top h-index in Canada for mathematics and statistics, and has >16000 citations and h=63. He received five best paper awards. He is Fellow of IEEE, Canadian Academy of Engineering and Academia Europaea. He received Humboldt Research Award in Germany and Royal Society Research Merit Award in UK. He is Tsinghua 1000 Plan Distinguished Professor.