A New Dependence Model for Heterogeneous Markov Modulated Poisson Processes

Fang Dong, Kui Wu, Venkatesh Srinivasan
Department of Computer Science, University of Victoria, B.C., Canada

***Abstract*—Markov Modulated Poisson Process (MMPP) has been extensively studied in random process theory and widely applied in various applications involving Poisson arrivals whose rate varies following a Markov process. The most general form of aggregated MMPP is the superposition of heterogeneous MMPPs (HeMMPP), in which each constituent MMPP has different parameters. Due to the generality of HeMMPP, studying its temporal dependence will benefit network traffic monitoring and traffic prediction. Modeling the temporal dependence of HeMMPP, however, is extremely hard because the total number of states in a HeMMPP increases exponentially with the number of states in constituent MMPPs. This paper tackles the above challenge with copula analysis. It not only presents a novel framework to capture the functional dependence structure of HeMMPP, but also provides a recursive algorithm to effectively calculate HeMMPP copula values. The theoretical analysis and the algorithms together offer a complete solution for modeling the temporal dependence of HeMMPP. Another contribution of the paper is the application of HeMMPP copula for traffic prediction.**

I. Introduction

Markov modulated Poisson process (MMPP) is a doubly stochastic Poisson process whose arrival rate is modulated by an irreducible continuous time Markov chain (CTMC) independent with the arrival process [6]. MMPP was first proposed by Yechiali and Naor to model non-homogeneous Poisson arrival process in queueing systems [17]. Specifically, the arrival process is a Poisson process with arrival rate λ_j whenever the CTMC is in state j. MMPP can effectively capture burst arrivals and sudden changes in arrivals since it can integrate significantly different rates into one model. This advantage makes MMPP a widely applied model for the arrival processes of network systems [13]. When multiple independent MMPP flows, each having a different set of parameters, arrive in a system, the total arrivals become the superposition of independent heterogeneous MMPPs (HeMMPP).

From the theoretical aspect, HeMMPP is the most general form of MMPP aggregate, because single MMPP and the superposition of independent homogeneous MMPPs (HoMMPP) are both special cases of HeMMPP. Therefore, the study of HeMMPP can benefit both real-world applications and theoretical performance analysis. *From the practical aspect*, HeMMPP traffic exists in many real-world applications. For instance, HeMMPP has been applied to generate self-similar traffic to Internet backbone [19]. HeMMPP can be used to capture multiple multimedia sources to a multimedia server because each multimedia traffic can be reasonably modeled with MMPP [16].

ISBN 978-3-903176-08-9

The temporal dependence of HeMMPP can help develop fitting methods that model real network traffic trace with HeMMPP [1] or predict the trend of arrivals [19]. Despite its importance, however, the dependence structure of HeMMPP is largely unknown. It can be shown that HeMMPP is still an MMPP [6], but analysing HeMMPP as one MMPP becomes intractable due to the exponential increase in the number of states [8]. For instance, modeling a HeMMPP consisting of two 20-state MMPPs is computationally difficult [8]. In other words, simply treating the superposition of MMPPs as one MMPP with a larger number of states will not work well. As such, we need to develop a different method to analyse the temporal dependence of HeMMPP.

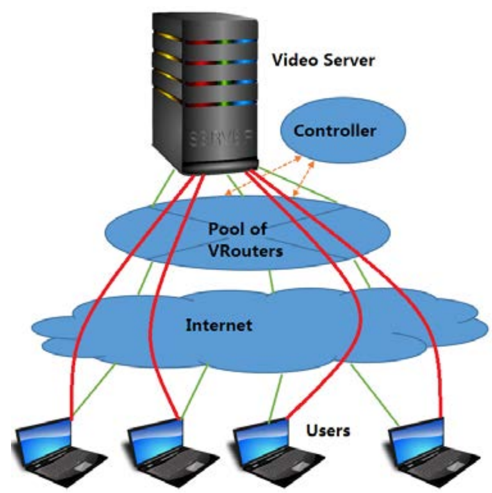

Fig. 1: Motivating example: a content delivery system relies on software-defined network and a pool of virtual network functions (e.g., VRouters) to adjust network bandwidth aligning with demand of the multimedia flows (denoted by red curves)

Astute readers may observe an alternative method to avoid the need of analysing HeMMPP: we may simply track and study each constituent MMPP, and in the context of traffic prediction, we predict the future arrivals of each constituent MMPP and aggregate them as the prediction of future HeMMPP arrivals. We call this alternative the *predict-individual* method. This method, however, poses extra burden in practice. In an example shown in Fig. 1, a content delivery system relies on software-defined networks and a pool of virtual network functions (e.g., VRouters) to adjust network bandwidth aligning with the multimedia flows, each labelled by a red curve and modeled as an MMPP [16]. Using the predict-individual method, the controller needs to keep record of arrivals of each MMPP and makes prediction on each MMPP. In contrast, with a HeMMPP model, the controller tracks arrivals of each MMPP *only during the HeMMPP modeling process*. Once the HeMMPP model is built, the controller does not need to track individual MMPP flows and instead uses the aggregated arrivals to predict future HeMMPP arrivals. Clearly, the HeMMPP model greatly simplifies the tasks of the controller. In addition, using HeMMPP model for prediction leads to more accurate results than using the predict-individual method, as shown in our later evaluation.

In the state-of-art temporal dependence model of HeMMPP, the covariance between number of arrivals in different time slots (called arrival counts) in constituent MMPP is derived asymptotically in [1]. The summation of asymptotic covariance of constituent MMPPs turns to be the approximate covariance between arrival counts in HeMMPP. Nevertheless, there is a large gap towards obtaining a complete temporal dependence model of HeMMPP for two reasons. First, most papers consider HeMMPP where its constituent MMPPs have only two states [1], [15]. The temporal dependence of HeMMPP, whose constituent MMPPs have a higher number of states, needs further study. Second, covariance or autocovariance is only capable of measuring *linear* dependence over time, but the network traffic traces may exhibit much more complex dependence than that. We are thus motivated to search for a functional dependence structure of HeMMPP, which carries richer and more complete information of dependence.

We tackle the problem with copula, an advanced dependence measure that links marginals into joint distribution. This paper analyses the functional dependence between arrival counts in HeMMPP and makes the following contributions:

1) It uses a new dependence measure, copula, to analyse the dependence structure of HeMMPP. The copula-based dependence reveals functional temporal dependence, which is more powerful than the commonly-used measures for linear dependence, covariance and correlation.
2) The copula-based analysis can effectively deal with the difficulty in modeling HeMMPP, where each constituent MMPP may have an arbitrary number of states.
3) It not only presents a recursive algorithm to compute the theoretical copula values of HeMMPP, but also adopts parametric copulas to model the temporal dependence of HeMMPP.
4) It demonstrates an application of HeMMPP dependence model in traffic prediction.

II. Preliminaries

A. Markov Modulated Poisson Process

We introduce the definition and key concepts of MMPP and HeMMPP.

Definition 1. *A **Markov-modulated Poisson Process** (MMPP) [6] is constructed by varying the arrival rate of a Poisson process according to an m-state irreducible continuous-time Markov chain (CTMC). In particular, when the Markov Chain is in state j, the arrivals follow a Poisson process of rate λ_j. Therefore, an MMPP can be parameterized by the Q matrix [14] of CTMC and the m Poisson arrival rates, $\Lambda = (\lambda_1, \ldots, \lambda_m)$.*

We thus denote an MMPP by parameters (Q, Λ).

Definition 2. ***Environment-stationarity** of an MMPP [6]: An MMPP (Q, Λ) is considered to be environment-stationary if its associated CTMC Q is stationary.*

For an environment-stationary MMPP, the stationary distribution of the states, $\Pi = (\pi_1, \ldots, \pi_m)$, is determined by solving the equation $\Pi Q = 0$.

Definition 3. ***HeMMPP**: An MMPP is called **HeMMPP** if it is a superposition of multiple independent heterogeneous MMPPs. The constituent MMPPs carry different parameters $({}_1Q, {}_1\Lambda), ..., ({}_rQ, {}_r\Lambda)..., ({}_lQ, {}_l\Lambda)$, where $({}_rQ, {}_r\Lambda)$ denotes the parameters of the r-th constituent MMPP.*

To distinguish regular MMPP with superposition of MMPPs, we use the term *single MMPP* to refer to an MMPP not created from superposition, the term *HeMMPP* to refer to aggregate MMPPs containing multiple single MMPPs, and each single MMPP in HeMMPP is called *constituent MMPP*. In this paper, we only consider stationary MMPPs, which means each constituent MMPP in HeMMPP is environment-stationary.

For ease of reference, the main notations used in the paper are listed in Tables I, II and III. Note that we consider HeMMPP with l number of constituent MMPPs. When $l = 1$, HeMMPP degrades to single MMPP, and in this case l can be omitted from notation. Thus A_i, M, C and $C_{i'}$ are notations for single MMPP.

Remark 1. *Due to the intricate composition of HeMMPP, a complicate and slightly unconventional notation system is necessary. We, however, adopt the following rules to make the notations easy to follow: the superscript denotes the number of constituent MMPPs, the right hand-side subscript denotes the time slot-related information (or random variables from the context), and the left hand-side subscript denotes the index of a specific constituent MMPP in consideration.*

B. Copulas

A copula is a function that links univariate marginals to their multivariate distribution. The definition of 2-copula is:

Definition 4. ***(Copula)** A 2-dimensional copula is a function C having the following properties [10]:*

1) *Its domain is $[0, 1] \times [0, 1]$;*
2) *C is 2-increasing, i.e., for every $u_1, u_2, v_1, v_2 \in [0, 1]$ and $u_1 \leq u_2, v_1 \leq v_2$, we have $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$.*
3) *$C(u, 0) = C(0, v) = 0$, $C(u, 1) = u$, $C(1, v) = v$, for every $u, v \in [0, 1]$.*

TABLE I: Parameters for Single MMPP

Notation	Explanation
Q	The transition rate matrix of associated CTMC
Λ	The vector of Poisson arrival rates
Π	The stationary distribution of asscoiated CTMC
$P(t)$	The transition matrix after time t of associated CTMC

TABLE II: Notations of HeMMPP

Notation	Explanation
A_i^l	The arrival count in i-th time slot of HeMMPP consisting of l number of consituent MMPPs
$M^l(x)$	The marginal distribution of A_i^l
$C^l(u,v)$	The copula between A_i^l and A_{i+1}^l
$\nabla C^l(u,v)$	The copula gradient of $C^l(u,v)$
$\mathbb{C}^l$	The copula matrix for $C^l(u,v)$
$\mathbb{R}^l$	The copula gradient matrix for $\nabla C^l(u,v)$
$C_{i'}^l(u,v)$	The copula between A_i^l and $A_{i+i'}^l$
$C_{i'}^l(u,v;\theta)$	The parametric copula between A_i^l and $A_{i+i'}^l$

Theorem 1. ***(Sklar's theorem)*** *[10] Let H be a joint distribution function with marginals F_X and F_Y, then there exists a copula C such that for for all x and y,*

$$H(x,y) = Pr(X \leq x, Y \leq y) = C(F_X(x), F_Y(y)).$$

Sklar's theorem is the core of the copula theory. First, it shows how the copula connects marginals with joint distribution. This property is especially useful since the joint distribution of random variables is hard to find directly in many applications [10]. In this situation, integration of a copula model and marginals makes it easy to understand the joint behaviour. Second, Sklar's theorem implies that copula, as a dependence measure, is entirely separated from both marginals and joint distribution.

The dependence in terms of copula is stable when the marginals changes functionally. This beautiful feature is formally stated in the following theorem:

Theorem 2. ***(The invariant property of copulas)*** *[10] Let X and Y be continuous random variables with copula C_{XY}. If α_1 and α_2 are strictly increasing functions on the range of X and the range of Y, respectively, then $C_{\alpha_1(X)\alpha_2(Y)} = C_{XY}$. In other words, C_{XY} is invariant under strictly increasing transformations of X and Y.*

There are mainly two methods to build a copula model. The first method is to construct a theoretical copula for the problem at hand. The inversion method belongs to this category.

Theorem 3. ***(Inversion method)*** *[10] Let H be a joint distribution function with marginals F_X and F_Y. Let F_X^{-1} and F_Y^{-1} be the inverse function of F_X and F_Y. Then the copula between X and Y can be constructed as*

$$C(u,v) = H(F_X^{-1}(u), F_Y^{-1}(v)) \quad \forall u, v,$$

such that

$$H(x,y) = C(F_X(x), F_Y(y)) \quad \forall x, y.$$

The other method to build a copula model is to fit real data into known parametric copulas. A large variety of parametric copulas are available for parametric copula modeling, for instance, Gaussian copula, Student's copula, Clayton copula, Frank copula, and Gumbel copula. Since the theoretical copula is not always easy to derive, parametric copula modeling has become popular in practice [7]. A copula model built with this method is also called an parametric copula.

TABLE III: Notations of the r-th Constituent MMPP

Notation	Explanation
$({}_rQ, {}_r\Lambda)$	The parameters of the r-th MMPP
${}_rA_i$	The arrival count in i-th time slot of r-th MMPP trace
${}_rM(x)$	The marginal distribution of ${}_rA_i$
${}_rp(x)$	The probability mass function of ${}_rA_i$
${}_rC(u,v)$	The copula between ${}_rA_i$ and ${}_rA_{i+1}$
$\nabla_r C(u,v)$	The copula gradient of ${}_rC(u,v)$
${}_rC_{i'}(u,v)$	The copula between ${}_rA_i$ and ${}_rA_{i+i'}$

Copula-based dependence is tightly associated with tail dependence measure. The tail dependence comprises upper tail dependence given by

$$\begin{aligned}\rho_t^+ &= \lim_{u\to 1} Pr(X > F_X^{-1}(u) | Y > F_Y^{-1}(u)) \\ &= \lim_{u\to 1} \frac{1 - 2u + C(u,u)}{1-u};\end{aligned} \quad (1)$$

and the lower tail dependence given by

$$\rho_t^- = \lim_{u\to 0} Pr(X < F^{-1}(u) | Y < F^{-1}(u)) = \lim_{u\to 0} \frac{C(u,u)}{u}. \quad (2)$$

Copula is promising for modeling temporal dependence of HeMMPP for the following reasons: copula can be constructed theoretically or by parametric copula modeling; copula is a functional dependence model and captures rounded dependence information beyond linear scope; the invariant property keeps copula structure stable when HeMMPP scales functionally. In this paper, we will study both theoretical copula (Section III) and parametric copula of HeMMPP (Section IV).

III. Theoretical Copula Analysis for HeMMPP

A. Theoretical Results for Single MMPP

A functional dependence model of single MMPP has been investigated in [4], from which some results are useful for our study of HeMMPP. To make this paper self-contained, we summarize these results below.

In the analysis of single MMPP with parameters (Q, Λ) [4], the time is divided into equal-sized small intervals, called time slots. The length of each time slot is denoted as Δ, which is short enough such that the state transition of MMPP within one time slot is negligible[1]. Denote the sequence of time slots as $I_1, I_2, \ldots, I_n$, and the number of arrivals in I_i as A_i. The sequence $\{A_i\}$ is also called arrival counts as introduced in Section I. Denote the transition matrix by $P(t) = [p_{j_1 j_2}(t)]$, where $p_{j_1 j_2}(t)$ is the probability that the CTMC switches from state j_1 to state j_2 after time t. The transition matrix $P(t)$ can be calculated with well-known methods such as those

[1]This assumption is justified since the arrival rate in one time slot is (approximately) stable when Δ is small.

introduced in Chapter 6.8 of [14]. The stationary distribution of CTMC is $\Pi = (\pi_1, \pi_2, ..., \pi_m)$. The marginal distribution function of A_i is given in Theorem 4, and the copula between A_i and A_{i+1} is given in Theorem 5. The multi-step copula between A_i and $A_{i+i'}$ is given in Theorem 6. These theoretical marginals and copulas can be applied to the constituent MMPP in HeMMPP.

Theorem 4. *[4] The marginal distribution of A_i is*

$$M(x) \equiv Pr(A_i \le x) = \sum_{j=1}^{m} \pi_j G_j(x) \tag{3}$$

where $G_j(x) = e^{-\lambda_j \Delta} \sum_{k=0}^{k=x} \frac{(\lambda_j \Delta)^k}{k!}$

Theorem 5. ***(MMPP copula)*** *[4] The copula of A_i and A_{i+1} can be calculated as:*

$$C(u, v) = \mathbb{G}(M^{-1}(u)) diag(\Pi) P(\Delta) \mathbb{G}(M^{-1}(v))^T, \tag{4}$$

where

- $\mathbb{G}(x) \equiv [G_1(x), \cdots, G_m(x)]$ *is a vector,*
- M^{-1} *is the inverse function of M defined by (3),*
- $diag(\Pi)$ *is a square diagonal matrix with the elements of vector Π on the main diagonal,*
- $\mathbb{G}(M^{-1}(v))^T$ *is the transpose of* $\mathbb{G}(M^{-1}(v))$.

Theorem 6. ***(Multi-step MMPP copula)*** *[4] The copula of A_i and $A_{i+i'}$ in a single MMPP can be calculated as:*

$$C_{i'}(u, v) = \mathbb{G}(M^{-1}(u)) diag(\Pi) P(i'\Delta) \mathbb{G}(M^{-1}(v))^T. \tag{5}$$

B. Theoretical Copula for HeMMPP

1) Theoretical Analysis for HeMMPP Copula: Since the constituent MMPPs in HeMMPP possess different parameters, we have to differentiate the constituent MMPPs by numbering them. We randomly select an order of constituent MMPPs, i.e., $({}_1Q, {}_1\Lambda), ({}_2Q, {}_2\Lambda), ..., ({}_rQ, {}_r\Lambda)..., ({}_lQ, {}_l\Lambda)$, where $({}_rQ, {}_r\Lambda)$ represents the parameters of the r-th constituent MMPP ($r = 1, 2, \cdots, l$). For constituent MMPPs, ${}_rA_i$, ${}_rM$, ${}_rp$, ${}_rC$, $\nabla_r C$ denote arrival counts, marginal CDF, marginal probability mass function (PMF), copula and copula gradient of r-th MMPP, respectively. In HeMMPP, A_i^l, C^l, M^l are notations of the superposition of the *first* l number of constituent MMPPs. Note that we introduce this ordering for ease of explanation, and the order will not influence our analytical results. The following theorems show how to analyse theoretical marginal and copula of HeMMPP.

Theorem 7. *The HeMMPP marginal distribution function has recursive relationship between M^l and $M^{l-1} (l \ge 2)$ as*

$$M^l(x) = \sum_{x'=0}^{x} M^{l-1}(x - x') * {}_lp(x'), \tag{6}$$

where ${}_lp$ is the probability mass function (PMF) of the arrival count from l-th MMPP, ${}_lp(x') = {}_lM(x') - {}_lM(x' - 1)$.

Proof. The key idea of the proof is to divide the arrivals from l number of MMPPs into the arrivals from the first $l - 1$ number of MMPPs plus the arrivals from the l-th MMPP, i.e., $A_i^l = A_i^{l-1} + {}_lA_i$. Thus, we have

$$\begin{aligned} M^l(x) = Pr(A_i^l \le x) &= \sum_{x'=0}^{x} Pr(A_i^l \le x | {}_lA_i = x') Pr({}_lA_i = x') \\ &= \sum_{x'=0}^{x} Pr(A_i^{l-1} \le x - x') Pr({}_lA_i = x') \\ &= \sum_{x'=0}^{x} M^{l-1}(x - x') * {}_lp(x') \end{aligned}$$

□

Theorem 8. *The HeMMPP copula has the recursive relationship between C^l and C^{l-1} as shown below:*

$$\begin{aligned} C^l(M^l(x), M^l(y)) = &\sum_{x'=0}^{x} \sum_{y'=0}^{y} \nabla_l C({}_lM(x'), {}_lM(y')) \\ &* C^{l-1}(M^{l-1}(x - x'), M^{l-1}(y - y')), \end{aligned} \tag{7}$$

where $\nabla_l C$ is the copula gradient of the l-th MMPP. The copula gradient of r-th MMPP ($r = 1, 2, \cdots, l$) is defined as

$$\begin{aligned} &\nabla_r C({}_rM(x'), {}_rM(y')) \\ \equiv &{}_rC({}_rM(x'), {}_rM(y')) + {}_rC({}_rM(x' - 1), {}_rM(y' - 1)) \\ &- {}_rC({}_rM(x'), {}_rM(y' - 1)) - {}_rC({}_rM(x' - 1), {}_rM(y')). \end{aligned} \tag{8}$$

Proof. The proof is also on the basis of $A_i^l = A_i^{l-1} + {}_lA_i$.

$$\begin{aligned} &C^l(M^l(x), M^l(y)) \\ =&Pr(A_i^l \le x, A_{i+1}^l \le y) \\ =&\sum_{x'=0}^{x} \sum_{y'=0}^{y} Pr(A_i^l \le x, A_{i+1}^l \le y | {}_lA_i = x', {}_lA_{i+1} = y') \\ &\quad * Pr({}_lA_i = x', {}_lA_{i+1} = y') \\ =&\sum_{x'=0}^{x} \sum_{y'=0}^{y} Pr(A_i^{l-1} \le x - x', A_{i+1}^{l-1} \le y - y') \\ &\quad * Pr({}_lA_i = x', {}_lA_{i+1} = y') \\ =&\sum_{x'=0}^{x} \sum_{y'=0}^{y} C^{l-1}(M^{l-1}(x - x'), M^{l-1}(y - y')) \\ &\quad * Pr({}_lA_i = x', {}_lA_{i+1} = y') \end{aligned}$$

Since the arrival counts follow discrete distribution and the domain is non-negative integers, for all non-negative integer x' and y', and $\forall r \in \{1, 2, \cdots, l\}$, we have

$$\begin{aligned} &Pr({}_rA_i = x', {}_rA_{i+1} = y') \\ =&Pr({}_rA_i \le x', {}_rA_{i+1} \le y') + Pr({}_rA_i \le x' - 1, {}_rA_{i+1} \le y' - 1) \\ &-Pr({}_rA_i \le x', {}_rA_{i+1} \le y' - 1) - Pr({}_rA_i \le x' - 1, {}_rA_{i+1} \le y') \\ =&{}_rC({}_rM(x'), {}_rM(y')) + {}_rC({}_rM(x' - 1), {}_rM(y' - 1)) \\ &-{}_rC({}_rM(x'), {}_rM(y' - 1)) - {}_rC({}_rM(x' - 1), {}_rM(y')) \\ =&\nabla_r C({}_rM(x'), {}_rM(y')). \end{aligned}$$

Therefore, by replacing $Pr({}_lA_i = x', {}_lA_{i+1} = y')$ with $\nabla_l C({}_lM(x'), {}_lM(y'))$, we can calculate the copula C^l as:

$$C^l(M^l(x), M^l(y)) = \sum_{x'=0}^{x}\sum_{y'=0}^{y} \nabla_l C({}_lM(x'), {}_lM(y')) * C^{l-1}(M^{l-1}(x-x'), M^{l-1}(y-y')).$$

□

Theorems 7 and 8 reveal the relationship between M^l and M^{l-1} and relationship between C^l and C^{l-1}. Even with these relationships, it is still hard to derive the closed-form HeMMPP copula. However, they are sufficient for developing recursive algorithms to numerically calculate HeMMPP copula, as introduced in the next section.

2) Recursive Algorithms for Calculating HeMMPP Copula: To design algorithms to calculate HeMMPP copula efficiently, we narrow down the interesting range of A_i^l from its infinite domain to finite range with an upper threshold $\hat{a}$. In other words, although the range of A_i^l is on the whole non-negative integer domain, we only need to compute the copula values $C^l(M^l(x), M^l(y))$ for $x < \hat{a}$ and $y < \hat{a}$. The selection of $\hat{a}$ is application dependent and can be set appropriately based on the trace data. Narrowing down the interesting range makes the computation feasible and still meets practical needs, because the arrival counts in real traffic flows always fall within a limited range.

On the interesting range $[0, \hat{a})$, we define seven matrices in Table IV. $\mathbb{M}^l$, $\mathbb{C}^l$ and $\mathbb{R}^l$ are for HeMMPP, and they represent HeMMPP marginal values, HeMMPP copula values and HeMMPP copula gradient values, respectively. For instance, the number in row x of matrix $\mathbb{M}^l$ represents the value of $M^l(x-1)$, *i.e.*, $\mathbb{M}^l_x \equiv M^l(x-1)$. For the constituent MMPPs, their values in PMF, marginal values, copula values and copula gradient values are represented by ${}_r\mathbb{P}$, ${}_r\mathbb{M}$, ${}_r\mathbb{C}$ and ${}_r\mathbb{R}$, respectively, as shown in Table IV. To emphasize the matrices' dimension, we mark dimensions on the bottom right, such as $[\mathbb{M}^l]_{\hat{a}}$ and $[\mathbb{C}^l]_{\hat{a}\times\hat{a}}$. We also use a notation $[\mathbb{C}^l]_{x\times y}$ to represent the submatrix of $[\mathbb{C}^l]_{\hat{a}\times\hat{a}}$ with its first x rows and first y columns.

With HeMMPP parameters $({}_1Q, {}_1\Lambda), ({}_2Q, {}_2\Lambda), ..., ({}_lQ, {}_l\Lambda)$ and a properly-set threshold value $\hat{a}$, we design Algorithm 1 (with the time complexity as $O(\hat{a} \times l)$) to calculate HeMMPP marginal matrix $[\mathbb{M}^l]_{\hat{a}}$ and Algorithm 2 (with the time complexity as $O(\hat{a} \times \hat{a} \times l)$) to calculate HeMMPP copula matrix $[\mathbb{C}^l]_{\hat{a}\times\hat{a}}$. The recursive procedure in Algorithm 1, **MARG, implements Theorem 7;** the recursive procedure in Algorithm 2, **CPA, implements Theorem 8**. With matrices $[\mathbb{M}^l]_{\hat{a}}$ and $[\mathbb{C}^l]_{\hat{a}\times\hat{a}}$ computed, the theoretical copula of HeMMPP is revealed in Theorem 9:

Theorem 9. ***(HeMMPP copula)*** *Given HeMMPP with marginal matrix $[\mathbb{M}^l]_{\hat{a}}$ and copula matrix $[\mathbb{C}^l]_{\hat{a}\times\hat{a}}$, its copula value of $C^l(u,v)$ can calculated by*

$$C^l(u,v) = \mathbb{C}^l_{(\mathrm{argmax}_x \mathbb{M}^l_x \le u)(\mathrm{argmax}_y \mathbb{M}^l_y \le v)} \quad (9)$$

for any u and v satisfying $u \le \mathbb{M}^l_{\hat{a}}$, $v \le \mathbb{M}^l_{\hat{a}}$.

Algorithm 1 An algorithm to compute marginal matrix $\mathbb{M}^l$

Require: the upper threshold $\hat{a}$, HeMMPP parameters $({}_1Q, {}_1\Lambda), ({}_2Q, {}_2\Lambda), ..., ({}_lQ, {}_l\Lambda)$,
Ensure: $[\mathbb{M}^l]_{\hat{a}}$
return MARG($[{}_1Q, ..., {}_lQ], [{}_1\Lambda, ..., {}_l\Lambda], \hat{a}$)

procedure MARG($[{}_1Q, ..., {}_lQ], [{}_1\Lambda, ..., {}_l\Lambda], \hat{a}$)
  $l \leftarrow$ the vector length of $[{}_1Q, ..., {}_lQ]$ or of $[{}_1\Lambda, ..., {}_l\Lambda]$
  // Base Case
  if $l == 1$ **then**
    $[\mathbb{M}^1]_{\hat{a}} \leftarrow$ compute with parameters ${}_1\Lambda$ and ${}_1Q$ based on Theorem 4
    return $[\mathbb{M}^1]_{\hat{a}}$
  end if
  // Inductive Step
  $[\mathbb{M}^{l-1}]_{\hat{a}} \leftarrow$ MARG($[{}_1Q, ..., {}_{l-1}Q]$, $[{}_1\Lambda, ..., {}_{l-1}\Lambda], \hat{a}$)
  $[{}_l\mathbb{M}]_{\hat{a}} \leftarrow$ compute with parameters ${}_l\Lambda$ and ${}_lQ$ based on Theorem 4
  $[{}_l\mathbb{P}]_{\hat{a}} \leftarrow$ compute from $[{}_l\mathbb{M}]_{\hat{a}}$ based on its definition
  for $x \leftarrow 1, \hat{a}$ **do**
    Rotate matrix $[{}_l\mathbb{P}]_x$ 180 degree clockwise as $[{}_l\mathbb{P}']_x$
    Calculate Hadamard product of $[\mathbb{M}^{l-1}]_x$ and $[{}_l\mathbb{P}']_x$ as $[\mathbb{T}]_x$
    $\mathbb{M}^l_x \leftarrow$ sum of all elements in matrix $[\mathbb{T}]_x$
  end for
  return $[\mathbb{M}^l]_{\hat{a}}$
end procedure

C. Multi-step Theoretical Copulas for HeMMPP

Theorem 10. ***(Multi-step HeMMPP copula)****. The copula of A_i^l and $A_{i+i'}^l$ in HeMMPP can be constructed by integrating the multi-step MMPP copula in Theorem 6 into the recursive method in Theorem 8. Specifically, all copulas C^l in the recursive method are replaced by multi-step copulas $C^l_{i'}$, and all constituent copulas ${}_rC$ are replaced by ${}_rC_{i'}$.*

IV. Parametric Copula modeling for HeMMPP

In this section, we construct parametric copulas for HeMMPP. The parametric copulas we investigated include the following three Archimedean copulas [10]:

1) Clayton copula ($\theta \in [-1, \infty) \setminus \{0\}$)

$$C(u, v; \theta) = [\max\{u^{-\theta} + v^{-\theta} - 1, 0\}]^{-1/\theta},$$

2) Frank copula ($\theta \in [-\infty, \infty) \setminus \{0\}$)

$$C(u, v; \theta) = -\tfrac{1}{\theta}\log[1 + \tfrac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{\exp(-\theta)-1}],$$

3) Gumbel copula ($\theta \in [1, \infty)$)

$$C(u, v; \theta) = \exp[-((-\log u)^\theta + (-\log v)^\theta)^{1/\theta}].$$

We investigate the above parametric copulas due to two reasons. First, they are all one-parameter copulas which make modelling easy. Second, they capture different types of tail dependence efficiently. The tail dependence features of the three copulas are distinct with each other: Clayton copula

TABLE IV: Definition of Matrices in HeMMPP

Matrix Denotation	Matrix name	Number in row x (and column y)
$[\mathbb{M}^l]_{\hat{a}}$	marginal matrix of HeMMPP	$\mathbb{M}^l_x \equiv M^l(x-1)$
$[\mathbb{C}^l]_{\hat{a}\times\hat{a}}$	copula matrix of HeMMPP	$\mathbb{C}^l_{xy} \equiv C^l(M^l(x-1), M^l(y-1))$
$[\mathbb{R}^l]_{\hat{a}\times\hat{a}}$	copula gradient matrix of HeMMPP	$\mathbb{R}^l_{xy} \equiv \nabla C^l(M^l(x-1), M^l(y-1))$
$[{}_r\mathbb{M}]_{\hat{a}}$	marginal matrix of r-th MMPP	${}_r\mathbb{M}_x \equiv {}_rM(x-1) = Pr({}_rA_i \leq x-1)$
$[{}_r\mathbb{P}]_{\hat{a}}$	PMF matrix of r-th MMPP	${}_r\mathbb{P}_x \equiv {}_rp(x-1) = {}_r\mathbb{M}_x - {}_r\mathbb{M}_{x-1}$
$[{}_r\mathbb{C}]_{\hat{a}\times\hat{a}}$	copula matrix of r-th MMPP	${}_r\mathbb{C}_{xy} \equiv {}_rC({}_rM(x-1), {}_rM(y-1))$
$[{}_r\mathbb{R}]_{\hat{a}\times\hat{a}}$	copula gradient matrix of r-th MMPP	${}_r\mathbb{R}_{xy} \equiv \nabla {}_rC({}_rM(x-1), {}_rM(y-1))$

Algorithm 2 An algorithm to compute copula matrix $\mathbb{C}^l$

Require: the upper threshold $\hat{a}$, HeMMPP parameters $({}_1Q, {}_1\Lambda), ({}_2Q, {}_2\Lambda), ..., ({}_lQ, {}_l\Lambda)$
Ensure: $[\mathbb{C}^l]_{\hat{a}\times\hat{a}}$

return CPA($[{}_1Q, ..., {}_lQ], [{}_1\Lambda, ..., {}_l\Lambda], \hat{a}$)

procedure CPA($[{}_1Q, ..., {}_lQ], [{}_1\Lambda, ..., {}_l\Lambda], \hat{a}$)
    $l \leftarrow$ the vector length of $[{}_1Q, ..., {}_lQ]$ or of $[{}_1\Lambda, ..., {}_l\Lambda]$
    // Base Case
    if $l == 1$ **then**
        $[\mathbb{C}^1]_{\hat{a}\times\hat{a}} \leftarrow$ compute with parameters ${}_1\Lambda$ and ${}_1Q$ based on Theorem 5
        return $[\mathbb{C}^1]_{\hat{a}\times\hat{a}}$
    end if
    // Inductive Step
    $[\mathbb{C}^{l-1}]_{\hat{a}\times\hat{a}} \leftarrow$ CPA($[{}_1Q, ..., {}_{l-1}Q]$, $[{}_1\Lambda, ..., {}_{l-1}\Lambda], \hat{a}$)
    $[{}_l\mathbb{C}]_{\hat{a}\times\hat{a}} \leftarrow$ compute with parameters ${}_l\Lambda$ and ${}_lQ$ based on Theorem 5
    $[{}_l\mathbb{R}]_{\hat{a}\times\hat{a}} \leftarrow$ compute from $[{}_l\mathbb{C}]_{\hat{a}\times\hat{a}}$ based on its definition
    for $x \leftarrow 1, \hat{a}$ **do**
        for $y \leftarrow 1, \hat{a}$ **do**
            Rotate matrix $[{}_l\mathbb{R}]_{x\times y}$ 180 degree clockwise to be $[{}_l\mathbb{R}']_{x\times y}$
            Calculate Hadamard product of $[\mathbb{C}^{l-1}]_{x\times y}$ and $[{}_l\mathbb{R}']_{x\times y}$ as $[\mathbb{T}]_{x\times y}$
            $\mathbb{C}^l_{xy} \leftarrow$ sum of all elements in matrix $[\mathbb{T}]_{x\times y}$
        end for
    end for
    return $[\mathbb{C}^l]_{\hat{a}\times\hat{a}}$
end procedure

models lower tail dependence; Gumbel copula models upper tail dependence; and Frank copula captures symmetric upper and lower tail dependence. Therefore, these three copulas are investigated as simple alternatives of theoretical copulas. To further improve copula fitting of HeMMPP, a mixture of these copulas or some other types of parametric copulas might be needed for modelling, which is, however, beyond the scope of this paper and a possible topic for extended research.

We follow three main steps to model the parametric copula:

1) Compute the tail dependence by definitions in Eq.(1) and Eq. (2).
2) Choose the parametric copula for HeMMPP modeling based on tail dependence:
   a) choose Clayton copula if $\rho_t^+ \approx 0$ and $\rho_t^- > 0$;
   b) choose Frank copula if $\rho_t^+ \approx \rho_t^-$;
   c) choose Gumbel copula if $\rho_t^+ > 0$ and $\rho_t^- \approx 0$.
3) Determine the value of the parameter θ for the chosen parametric copula. The parameter θ is learnt by fitting the HeMMPP trace into the chosen copula with the maximum likelihood estimation method. Specifically, θ of parametric HeMMPP copula $C^l_{i'}(u, v; \theta)$ is determined by fitting the sample pairs of (A^l_i, $A^l_{i+i'}$) of trace.

V. Application: Traffic Prediction Based on HeMMPP Copula

So far, we have shown how the full temporal dependence structure of HeMMPP can be captured with copulas. A deep understanding of the temporal dependence structure of HeMMPP can benefit many applications, e.g., dynamic resource provisioning, and self-similar traffic modeling. Another obvious application is to predict future traffic based on the temporal dependence in arrivals. This section illustrates a method to achieve this goal.

The problem of traffic prediction could be in different forms. In this paper, we focus on estimating the future arrival count $A^l_{i+i'}$ based on the current observation of arrival count A^l_i. The prediction is made by maximizing the conditional probability $Pr(A^l_{i+i'}|A^l_i)$. When $i' = 1$, the prediction is made one-step forward; when $i' > 1$, the prediction is made multi-step forward. In this section, we introduce the prediction methods with both theoretical and parametric HeMMPP copula.

Prediction based on theoretical HeMMPP copula is made according to Theorem 11:

Theorem 11. *(1) Consider a HeMMPP having theoretical copula C^l between A^l_i and A^l_{i+1}. If $A^l_i = x$ is the current observation from the arrival process and if the prediction is made by maximizing the conditional probability $Pr(A^l_{i+1}|A^l_i)$, the predicted arrival count $\hat{A}^l_{i+1}$ is:*

$$\hat{A}^l_{i+1} = \operatorname*{argmax}_y \nabla C^l(M^l(x), M^l(y)). \tag{10}$$

(2) Consider a HeMMPP having theoretical copula $C_{i'}^l$ between A_i^l and $A_{i+i'}^l$. If $A_i^l = x$ is the current observation from the arrival process and if the prediction is made by maximizing the conditional probability $Pr(A_{i+i'}^l | A_i^l)$, the predicted arrival count $\hat{A}_{i+i'}^l$ is:

$$\hat{A}_{i+i'}^l = \operatorname*{argmax}_y \nabla C_{i'}^l(M^l(x), M^l(y)). \tag{11}$$

∇C^l and $\nabla C_{i'}^l$ are copula gradients defined in the same way as in Eq.(8) by replacing ${}_rC$ with C^l or $C_{i'}^l$ and ${}_rM$ with M^l.

Proof. We only prove part (1), because part (2) can be proved in the same way. Since the prediction is made by maximizing the conditional probability $Pr(A_{i+1}^l | A_i^l)$, we have

$$\begin{aligned}\hat{A}_{i+1} &= \operatorname*{argmax}_y Pr(A_{i+1}^l = y | A_i^l = x)\\ &= \operatorname*{argmax}_y \frac{Pr(A_i^l = x, A_{i+1}^l = y)}{Pr(A_i^l = x)}\\ &= \operatorname*{argmax}_y \frac{\nabla C^l(M^l(x), M^l(y))}{Pr(A_i^l = x)}\\ &= \operatorname*{argmax}_y \nabla C^l(M^l(x), M^l(y))\end{aligned}$$

□

Prediction based on parametric HeMMPP copula is made in the similar way of theoretical HeMMPP copula, as shown in Theorem 12. The proof is omitted because the proof idea is the same as that of Theorem 11.

Theorem 12. *Consider a HeMMPP having parametric copula $C_{i'}^l(u, v; \theta)$ between A_i^l and $A_{i+i'}^l$. If $A_i^l = x$ is the current observation from the arrival process and if the prediction is made by maximizing the conditional probability $Pr(A_{i+i'}^l | A_i^l)$, the predicted arrival count $\hat{A}_{i+i'}^l$ is:*

$$\hat{A}_{i+i'}^l = \operatorname*{argmax}_y \nabla C_{i'}^l(M^l(x), M^l(y); \theta), \tag{12}$$

where $\nabla C_{i'}^l(M^l(x), M^l(y); \theta)$ is copula gradient defined in the same way as in Eq. (8) by using copula $C_{i'}^l(M^l(x), M^l(y); \theta)$ and marginal M^l.

VI. Experimental Evaluation

A. Evaluation Methods

We have proposed both theoretical copula modeling and parametric copula modeling for traffic prediction in Section V. To evaluate the new model, we implement another prediction model, linear predictive coding (LPC(1)), for comparison. LPC(1) will be constructed from trace data. With LPC(1), the multi-step prediction from A_i^l is made as

$$\hat{A}_{i+1}^l = \gamma A_i^l, \quad \hat{A}_{i+2}^l = \gamma \hat{A}_{i+1}^l, \quad \cdots, \quad \hat{A}_{i+i'}^l = \gamma \hat{A}_{i+i'-1}^l,$$

where γ is the parameter of LPC(1) model.

LPC(1) model predicts data based on the dependence information in terms of autocorrelation. Thus it is set as the benchmark predictor to show *how functional dependence modeling with copulas improves over linear dependence.* Note that the first order of LPC model is used here for a fair comparison: our copula-based prediction model is first order in the sense that only dependence between two arrival counts is considered each step. It has been shown that copula models outperform AR(1) model in MMPP traffic prediction [4]. Due to space limit, however, we omit this comparison.

We also implement and compare the *predict-individual* method, where we predict the future arrivals of each constituent MMPP separately and aggregate them as the prediction of future HeMMPP arrivals.

When applying any of the prediction models on a traffic trace, the trace is divided into two parts, the training set and the testing set. The training set comes from the first certain percentage data of the trace, and the rest of the trace constitutes the testing set. The prediction accuracy is measured by root-mean-square error (RMSE) across the testing set:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{A}_i^l - A_i^l)^2}, \tag{13}$$

where A_i^l is arrival counts from testing set at timeslot i, $\hat{A}_i$ denotes the corresponding predicted value, and n is the total number of time slots in the testing period. For a prediction model, its performance is measured by its average RMSE (aRMSE) on a trace with different training percentages. The performance improvement ratio (IMP RATIO) over benchmark model (LPC(1)) are defined as Eq.(14).

$$\text{IMP RATIO} = \frac{\text{aRMSE}_{\text{benchmark}} - \text{aRMSE}}{\text{aRMSE}_{\text{benchmark}}} * 100\%. \tag{14}$$

B. Evaluation Results

We evaluate the benefit of using HeMMPP dependence model for traffic prediction with synthetic data. We generate a HeMMPP trace using two MMPP models, each obtained by fitting the model to real-world trace. For this purpose, we follow the work [4] and use the same Bellcore traces[2], that record millions of packet arrivals on an Ethernet at Bellcore Morristown Research and Engineering facility. The traces are well known in network traffic modeling, and many papers have shown that Bellcore traces are well characterized by MMPP [1], [9]. We choose two of these traces to determine reasonable parameters for simulation of synthetic HeMMPP trace. With the fitting algorithm in [8], BCpAug89 trace is well fitted into a 12 state MMPP [4] with parameters (Q_A, Λ_A) as shown in Eq. (15); BCpOct89 trace is fitted into a 13 state MMPP with parameters (Q_O, Λ_O) listed in Eq. (16). We generate synthetic HeMMPP data consisting of two MMPP traces, by simulating each MMPP trace using the learnt parameters and aggregating the two MMPP traces.

The length of time slots is set as $\Delta = 1$ second. A_i^l denotes the number arrival of the aggregate trace in i-th second. We will study one-step dependence between A_i^l and A_{i+1}^l and two-step dependence between A_i^l and A_{i+2}^l, and conduct one-step prediction and two-step prediction accordingly.

[2]available from the website http:// ita.ee.lbl.gov/html/contrib/BC.html

$$Q_A = \begin{pmatrix} -0.857 & 0.286 & 0.429 & 0.143 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.067 & -0.900 & 0.267 & 0.233 & 0.233 & 0.067 & 0.033 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.023 & 0.078 & -0.836 & 0.336 & 0.203 & 0.102 & 0.078 & 0.000 & 0.016 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.026 & 0.140 & -0.720 & 0.274 & 0.153 & 0.085 & 0.029 & 0.007 & 0.007 & 0.000 & 0.000 \\ 0.002 & 0.008 & 0.051 & 0.173 & -0.650 & 0.244 & 0.122 & 0.041 & 0.006 & 0.002 & 0.002 & 0.000 \\ 0.000 & 0.001 & 0.027 & 0.073 & 0.173 & -0.696 & 0.303 & 0.094 & 0.014 & 0.009 & 0.001 & 0.000 \\ 0.000 & 0.001 & 0.004 & 0.019 & 0.099 & 0.233 & -0.616 & 0.200 & 0.048 & 0.012 & 0.001 & 0.000 \\ 0.000 & 0.000 & 0.008 & 0.023 & 0.049 & 0.184 & 0.409 & -0.775 & 0.084 & 0.015 & 0.003 & 0.000 \\ 0.000 & 0.000 & 0.008 & 0.015 & 0.015 & 0.120 & 0.301 & 0.218 & -0.805 & 0.113 & 0.015 & 0.000 \\ 0.000 & 0.020 & 0.000 & 0.000 & 0.059 & 0.059 & 0.235 & 0.078 & 0.275 & -0.824 & 0.098 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.077 & 0.231 & 0.231 & 0.154 & 0.077 & -0.846 & 0.077 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 1.000 & 0.000 & -1.000 \end{pmatrix}, \quad (15)$$

$$\Lambda_A = \big(782.069, 674.207, 574.345, 482.483, 398.621, 322.759, 254.897, 195.035, 143.173, 99.311, 63.449, 35.587\big).$$

$$Q_O = \begin{pmatrix} -1.00 & 0.75 & 0.00 & 0.25 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.04 & -0.64 & 0.26 & 0.25 & 0.06 & 0.02 & 0.00 & 0.02 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.13 & -0.72 & 0.34 & 0.16 & 0.03 & 0.03 & 0.02 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.01 & 0.06 & 0.12 & -0.68 & 0.31 & 0.13 & 0.04 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.10 & 0.25 & -0.74 & 0.20 & 0.11 & 0.06 & 0.01 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.04 & 0.09 & 0.23 & -0.71 & 0.20 & 0.10 & 0.03 & 0.02 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.03 & 0.06 & 0.31 & -0.68 & 0.16 & 0.08 & 0.02 & 0.01 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.01 & 0.02 & 0.04 & 0.19 & 0.34 & -0.81 & 0.16 & 0.05 & 0.01 & 0.01 & 0.01 \\ 0.00 & 0.00 & 0.00 & 0.01 & 0.04 & 0.09 & 0.23 & 0.29 & -0.83 & 0.14 & 0.04 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.03 & 0.02 & 0.07 & 0.22 & 0.28 & -0.80 & 0.13 & 0.05 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.13 & 0.21 & 0.33 & -0.71 & 0.04 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.17 & 0.50 & 0.17 & -0.83 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 0.00 & 0.00 & 0.00 & -1.00 \end{pmatrix}, \quad (16)$$

$$\Lambda_O = \big(1125.89, 995.67, 873.46, 759.24, 653.02, 554.81, 464.59, 382.37, 308.15, 241.94, 183.72, 133.50, 91.28\big).$$

1) One-step Prediction on HeMMPP trace: Theoretical HeMMPP copula, parametric HeMMPP copula and LPC(1) model are constructed for one-step prediction. Theoretical HeMMPP copula is computed with Algorithms 1 and 2. Based on the observations of the HeMMPP trace, the threshold for marginal and copula matrix computation is chosen as $\hat{a} = 1500$. The probability that the arrival count A_i^l exceeds the threshold is less than 0.01, *i.e.*, $Pr(A_i^l > \hat{a}) < 0.01$, indicating that there are very few observations appearing beyond the chosen threshold. Fig. 2 shows the contour of theoretical one-step HeMMPP copula calculated with parameters $(Q_A, \Lambda_A, Q_O, \Lambda_O)$.

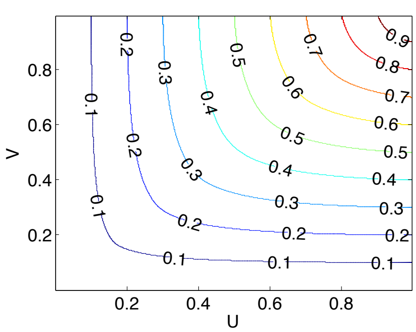

Fig. 2: The contour of theoretical one-step HeMMPP copula

Parametric HeMMPP copula is constructed through the modeling steps illustrated in Section IV. As the upper tail dependence is close to lower tail dependence ($\rho_t^+ = 0.3212$, $\rho_t^- = 0.2228$), Frank copula is chosen and its parameter is determined by fitting the training set of trace. Similarly, the parameter of LPC(1) model is determined according to the training set of the HeMMPP trace.

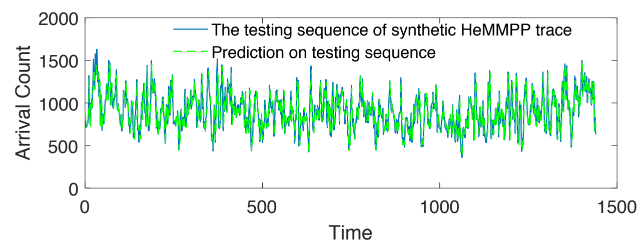

Fig. 3: One-step prediction with theoretical HeMMPP copula.

TABLE V: One-Step Prediction RMSE on the HeMMPP trace with Different Training Percentage.

Training Percentage	Theoretical Copula	Parametric Copula	Predict Individual	LPC(1)
50%	129.8456	129.9517	133.2521	190.6223
60%	129.1777	127.7076	132.0675	190.2156
70%	130.7527	129.9172	134.2631	193.6528
80%	129.4203	128.5920	133.5243	191.6844
90%	124.9783	125.2756	128.6330	188.7742
aRMSE	128.8349	128.2888	132.3480	190.9899
IMP RATIO	32.54%	32.83%	30.70%	—

With different percentage of trace data for training, three models are constructed accordingly and applied for one-step prediction. Fig. 3 shows the prediction with theoretical HeMMPP copula on the test set of the last 20% arrival counts. The detailed prediction errors are shown in terms of RMSE in Table V. The smaller value of RMSE represents the more accurate prediction. aRMSE shows the average performance over different training percentage (from 50% to 90%). IMP RATIO shows that **copula-based predictions, including theoretical**

copula model and parametric copula model, have more than 30% improvement over the LPC(1) model, showing the advantage of functional dependence modeling (such as copulas) over linear dependence measurement (such as autocorrelation). In addition, we can see that HeMMPP-based prediction works better than the predict-individual method. This is because any prediction includes errors and the predict-individual method may aggregate the errors from predictions in individual MMPP. This phenomenon is more obvious in the two-step prediction results shown in Table VI.

2) Two-step Prediction on HeMMPP trace: Two-step prediction is also performed to evaluate multi-step dependence modeling. Similar to the previous section, theoretical copula, parametric copula and LPC(1) models are constructed for two-step prediction. Table VI compares the two-step prediction performance of copulas with LPC(1) model. **Our copula models have a great improvement ratio (nearly 30%) over LPC(1) model regarding the two-step predictions.**

TABLE VI: Two-Step Prediction RMSE on the HeMMPP trace with Different Training Percentage.

Training Percentage	Theoretical Copula	Parametric Copula	Predict Individual	LPC(1)
50%	174.8214	175.6217	195.0492	246.4333
60%	174.4992	174.4648	192.8411	245.1160
70%	176.9906	176.0216	197.3482	252.5563
80%	174.9567	174.0762	194.7852	247.4094
90%	170.2201	169.0609	194.4664	246.7953
aRMSE	174.2976	173.8490	194.8980	247.6621
IMP RATIO	29.62%	29.80%	21.30%	—

VII. Related Work

The research related to our work applies covariance to model the temporal dependence among single MMPP or HeMMPP. The covariance between arrival counts in single MMPP is derived in [11]. The closed-form covariance between arrival counts in two-state MMPP is given in [1]. The work closest to ours is [4], where temporal dependence in single MMPP and in superposition of homogeneous MMPPs has been investigated. Nevertheless, HeMMPP is more complicated and more general, and its temporal dependence has not been studied in [4].

In the case of HeMMPP, the constituent MMPPs can be combined into one MMPP with formula given in [6]. Even though some efforts have made to reduce the number of states to approximate HeMMPP [8], [18], very few works consider HeMMPP as one MMPP for calculation of covariance. In [1], asymptotic covariance of two-state MMPPs is summed to get an approximate covariance of the superposition of MMPPs. Our work is different from the above work as we build the functional temporal dependence of arrival counts in HeMMPP with copula.

Copula models have been broadly used in the domain of financial analysis, for multivariate dependence modeling [3] as well as for time series modeling [12]. It has attracted attention in the domain of networks in recent years [2], [4], [5].

VIII. Conclusion

With copula analysis, this paper is the first that theoretically derives the intricate temporal dependence structure in HeMMPP. It not only presents a complete solution for modeling functional dependence in HeMMPP, but also introduces parametric copulas as fast approximation of theoretical copulas. Using the theoretical and parametric copulas, we show the value of our research in an example application, i.e., traffic prediction. While the study of MMPP has a long history and the topic of MMPP might not be trending, the novel theoretical results and the new algorithms developed in this paper will benefit a broad class of current and future network applications involving MMPP traffic flows.

References

[1] A. T. Andersen and B. F. Nielsen. A markovian approach for modeling packet traffic with long-range dependence. *IEEE Journal on Selected Areas in Communications*, 16(5):719–732, 1998.

[2] K. Avrachenkov, N. M. Markovich, and J. K. Sreedharan. Distribution and dependence of extremes in network sampling processes. *Computational Social Networks*, 2(1):1, 2015.

[3] M. Beil. *Modeling dependencies among financial asset returns using copulas*. PhD thesis, Technische Universität München, 2013.

[4] F. Dong, K. Wu, and V. Srinivasan. Copula analysis of temporal dependence structure in markov modulated poisson process and its applications. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, 2(4), 2017.

[5] F. Dong, K. Wu, and S. Venkatesh. Copula analysis for statistical network calculus. In *Proceedings of INFOCOM*, pages 1535–1543, Hong Kong, 2015. IEEE.

[6] W. Fischer and K. Meier-Hellstern. The markov-modulated poisson process (mmpp) cookbook. *Performance Evaluation*, 18(2):149–171, 1993.

[7] C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2):199–213, 2009.

[8] D. P. Heyman and D. Lucantoni. Modeling multiple ip traffic streams with rate limits. *IEEE/ACM Transactions on Networking*, 11(6):948–958, 2003.

[9] L. Muscariello, M. Mellia, M. Meo, M. A. Marsan, and R. L. Cigno. Markov models of internet traffic and a new hierarchical mmpp model. *Computer Communications*, 28(16):1835–1851, 2005.

[10] R. B. Nelsen. *An introduction to copulas*. Springer, New York, 2006.

[11] M. F. Neuts. *Structured Stochastic Matrices of M/G/1 Type and Their Applications*. Taylor & Francis, New York, USA, 1989.

[12] A. Patton. Copula methods for forecasting multivariate time series. *Handbook of economic forecasting*, 2:899–960, 2012.

[13] A. Rajabi and J. W. Wong. Provisioning of computing resources for web applications under time-varying traffic. In *2014 IEEE 22nd International Symposium on MASCOTS*, pages 152–157, Paris, France, 2014. IEEE.

[14] S. M. Ross. *Introduction to probability models*. Academic Press, Burlington, 2003.

[15] P. Salvador, R. Valadas, and A. Pacheco. Multiscale fitting procedure using markov modulated poisson processes. *Telecommunication Systems*, 23(1-2):123–148, 2003.

[16] S. Shah-Heydari and T. Le-Ngoc. Mmpp models for multimedia traffic. *Telecommunication Systems*, 15(3-4):273–293, 2000.

[17] U. Yechiali and P. Naor. Queuing problems with heterogeneous arrivals and service. *Operations Research*, 19(3):722–734, 1971.

[18] M. Yu and M. Zhou. A model reduction method for traffic described by mmpp with unknown rate limit. *IEEE Communications Letters*, 10(4):302–304, 2006.

[19] M. Yuksel, B. Sikdar, K. Vastola, and B. Szymanski. Workload generation for ns simulations of wide area networks and the internet. In *Proceedings of Communication Networks and Distributed Systems Modeling and Simulation Conference*, pages 93–98, 2000.