Abstract—This work studies end-to-end light-path assessment across network domains with incomplete management information. The focus is on investigating (a) whether the assessment can be made using domain-centric average information, (b) how to estimate such information using distributed algorithms which only exchange several parameters among neighboring domains, and (c) how informative these parameters are. We apply probabilistic graphical models and Mean-Field approximations to study these issues. We derive an algorithm for distributed information-estimation, a simple decision rule for path availability using the local and the average algorithm for distributed information-estimation, a simple decision rule for path availability using the local and the average information, and performance bounds for path assessment in terms of blocking probabilities.

Keywords—Network management, optical networks, probabilistic graphical models, machine learning

I. INTRODUCTION

A. Motivation

A key characteristic of optical networks is the use of light-paths for supporting end-to-end connections. Light-path assessment determines the availability of resources on a path across network domains, and thus plays a basic and important role in such network management and control functions as path selection, scheduling and wavelength routing.

Consider one path as an example where the assessment is done at the first domain, e.g. domain A in Figure 1. Assume that the path can support two types of traffic: inter-domain traffic that can originate at any subnet and traverse multiple domains; and intra-domain traffic that can originate and terminate within one domain [1]. The path assessment is done at the head domain (i.e., domain A). The information available there includes complete wavelength usage within the domain and incomplete information from the domains downstream.

Management information and performance are closely related quantities. For example, when complete information on wavelength usage is available from domains downstream, there would be no assessment error. If the information is incomplete, assessments could be erroneous. The more information and the exchange, the more accurate and the easier the assessment. Meanwhile, the more detailed the information used for assessment, the less efforts are needed in information-estimation. However, the cost grows when the amount of management information and the exchange increases. Hence it is important to provide a joint study on the performance of path assessment and the amount of management information.

B. Summary of Approaches and Results

1) Approaches

This work focuses on average (and aggregated) domain-centric information and the exchange. Aggregated information can be instantaneous and random such as the number of wavelengths used by inter-domain calls at subnets. Average information can be deterministic such as statistics of wavelength usage at domains. The average (and aggregated) information we study corresponds to a certain parameters that can be estimated and transmitted across domains. These parameters, if chosen properly, can conceal sensitive
information within subnets while being representative of wavelength usage at a domain. The problem is whether path assessment can be done based on the complete local information within a domain, and average domain-centric information from other subnets.

Path assessment considered in this work includes two tasks: (a) estimation of domain-centric average information, and (b) decision on path availability using the information obtained. The estimation process involves obtaining the parameters. There are three commonly-used schemes for information estimation: centralized, distributed and decentralized. A centralized scheme obtains and/or performs information estimation at a centralized location, e.g. domain A. A distributed scheme as shown in Figure 2 estimates domain-centric information locally at individual subnets by communicating with neighbors. Such distributed estimation is desirable where domains exchange information locally through self-coordination. A decentralized scheme is a special case of the distributed approach, where there is no information exchange among domains, and each subnet performs estimation based on only local information. Decision on path availability is centralized for three schemes as shown in Figure 3 using local information from the head domain and the average information from subnets downstream.

![Figure 2. Distributed estimation of domain-centric information.](image)

The performance is defined for path assessment as the probability of an incorrect decision [1]. Such a performance is path-centric and thus global; whereas the distributed estimation of information is domain-centric and thus local. Estimation and decision are then formulated jointly so that the domain-centric information is estimated by optimizing the performance of path assessment. In this work, we focus on distributed information estimation.

Two approaches we apply to deriving distributed algorithms for information estimation are probabilistic graphical models and Mean-Field approximation [2][3]. Probabilistic graphical models have been studied widely in machine learning and information theory [2][3], applied in many areas [3][4], and recently in networking [31-33].

A key role of probabilistic graphical models in this work is to provide an explicit representation of spatial dependence of aggregated wavelength usage at domains. Such a representation provides a spatial model of domain variables. Some emerging applications of the graphical models take a top-down approach by assuming the structure of a model [6]. This work chooses a bottom-up approach by first considering commonly-used traffic patterns at WDM layer, and then examining the graphical structure. This is because the structure of a probabilistic graph is determined by the spatial dependence among domains which results from traffic patterns. The structure of the graph shows which and how many neighboring domains to exchange information with when estimating the average information. Distributions of inter- and intra-domain traffic then determine specific parameters to be exchanged among domains.

For example, when inter-domain calls follow a commonly-used traffic pattern whose probability of traversing multiple domains decreases exponentially [1], the graphical model exhibits a spatial Markov dependence. A Mean-Field solution can then be derived based on the graph, resulting in domain-centric average information through communication only among the nearest neighbors.

2) Results:

(a) We formulate path assessment as a decision problem using local observations of wavelength usage in the head domain and average information from other subnets. The optimal assessment corresponds to the Bayes rule, resulting in the smallest decision error [24]. Here, global information on wavelength usage is needed from the entire path.

(b) We first describe an optimal distributed algorithm that minimizes the probability of erroneous decisions. As the optimal distributed algorithm is computationally costly for a large network, we develop a sub-optimal Mean-Field solution to decompose the blocking probability of the path into independent factors of the domains. The Mean-Field equations are coupled due to the spatial dependence shown by the graphical structure. A distributed algorithm solves the Mean-Field equations iteratively by exchanging several parameters among neighboring domains. The Mean-Field approximation in this work is novel as it is developed in the Fourier domain for accurate approximation.

(c) In the estimation stage, the information sent by domain \( k \) (1 ≤ \( k \) ≤ \( L \)) to its neighbors is a set of parameters, \( V_k = \{v_k, \log a_k, c_{ik}, d_{ik}, \gamma_k \} \), pertaining to this domain. \( c_{ik} \) and \( d_{ik} \) correspond to the second moments of inter-domain calls, \( \gamma_k \) is a Mean-Field adjustment to the average value of the wavelength usage of inter-domain calls, \( a_k \) is the probability of a wavelength occupied by an intra-domain call, and \( V_k \) is related to the mean-field decomposition of the non-blocking probability of domain \( k \).
In the decision stage, the average information related to \( \gamma_i \) and \( \nu_k \) for \( 1 \leq k \leq L \) is sent to the head domain for path assessment.

(d) A simple decision rule is used by the head domain for light-path assessment using the estimated parameters as well as the complete local wavelength usage.

(e) The performance of the simple decision rule is obtained as the probability \( P_e \) of erroneous assessment,

\[ P_e \leq P_p \leq \min(P_e,1-P_e) + 3\Delta, \]

where \( P_e \) is the Bayes error, \( \min(P_e,1-P_e) \) is an upper bound achievable by the best centralized assessment, \( P_p \) is the blocking probability of the path, and \( \Delta \) is the average error due to mean field approximations. Comparisons are made on performance and the amount of information among the centralized, distributed, and decentralized schemes.

II. RELATED WORK

Network management with minimal and partial information is an important issue, especially in the context of inter-domain management, where usually it is not feasible to use complete information. In contrast to using complete information, one method is to manage network domains as separate entities [6-8]. The corresponding performance (i.e., the correctness of an assessment) can be poor due to lack of information. For instance, in inter-domain routing, upon the arrival of a connection request, the source network domain may first select the next network domain in the inter-domain route. Then each network domain along the inter-domain route checks wavelength availability within its own domain and select next domain in the inter-domain route in a sequential way. In this case, each network domain makes decentralized decisions for inter-domain routing, which results in two potential problems due to the missing management information: (a) Inter-domain route may not be optimal based on different metrics, e.g., wavelength availability and resource utilization, and (b) sometime, a call that is accepted by network domains at the beginning of the inter-domain route may be rejected by network domains downstream. An intermediate approach is to use partial information-exchange among network domains. For instance, a framework of inter-domain routing is proposed in [14] based on OSPF, where border nodes at domain boundaries form an OSPF routing area.

There are also a lot of efforts in standardization bodies on inter-domain management in optical networks [7]. The GMPLS framework considers three different models based on the sharing of topological and other network management information, e.g., on routing and link state, between the carrier and the client networks [7][10][23]. This can be complete information sharing (overlay model), customized information sharing (augmented model), and no information sharing at all (overlay model). The OIF proposes the (User-Network-Interface) UNI and (Network-Network-Interface) NNI framework [7]. The UNI specification corresponds to the overlay model in GMPLS framework, and the standardization for NNI is still under way. Generally the information sharing should be consistent with the operational needs and constraints imposed by both carriers and customers [7]. Very often, scalability and/or security of management information is one of the major issues, which make it necessary to use minimal or partial management information.

The idea of using partial information has also been investigated in several research field in optical networks such as network survivability [9][11][12], and wavelength routing [13]. There the focus is mainly on developing management approaches, making use of partial information. It is also suggested to use wavelength converters to reduce the amount management information [16][17].

Prior investigations in other related areas are also beneficial to this research. In particular, inaccurate or aggregated information is first investigated in the context of QoS routing for IP networks [18], where two examples are bandwidth and topology aggregation [18][19][20]. Recent work [34] considers a joint optimization of transport-layer and routing protocols. There, an interesting idea is to pass a certain parameters across layers.

The problem of light-path assessment can be considered as call control under the framework of Automatic Switched Optical Network [7]. Similar problem of route control has been investigated for peer-to-peer network [21], where the goal is to find the path with a desired performance and fault tolerance.

Key issues which remain open are quantifications of incomplete information versus performance, and development of distributed algorithms for information estimation.

III. PROBLEM DESCRIPTION

Consider light-path assessment for a single path, and assume that the first domain determines the path availability upon a request of an end-end connection. The path results in a bus network as shown in Figure 1. Wavelength converters are assumed to be located only on the domain boundaries but not within domains. Complete local information is assumed available to a managed domain \( i \) which includes its topology as the number of hops \( H_i \) on the path in subnet \( i \), the total number of wavelengths \( F_i \) at each hop, the number of wavelengths used by inter-domain calls \( X_i \), and that used by intra-domain calls \( Y_i \). Assume there are \( L \) domains.

We formulate path assessment as a decision problem [1]. Let \( \Omega \) be a binary variable, where \( \Omega = 1 \) corresponds to the availability of a light-path; and \( \Omega = 0 \), otherwise. Let \( I_i \) be an indicator variable, \( I_i = 1 \), if domain \( i \) has a wavelength available to support an end-end call, and \( I_i = 0 \), otherwise. Let \( I = \{I_1,\ldots,I_L\} \). An assessment is incorrect if a decision differs from the ground truth. The probability of an erroneous assessment is

\[ P_e = E[P(\Omega = 1, I) = 0 \mid X_i, Y_i) + P(\Omega = 0, I = 1 \mid X_i, Y_i)], \]
where $E[ ]$ is the expectation on local observations $X_i$ and $Y_i$. The performance of path assessment is measured by $P_e$.

Let $V_i$ be the average information associated with subnet $i$. The problem of information estimation is to obtain $V_i$ which minimizes the probability of an assessment error for $1 \leq i \leq L$. As inter-domain traffic introduces dependence among domain, estimating $V_i$ should be coupled with other domains, resulting in inter-domain information-exchange. Distributed estimation is desirable as domains only exchange information with their neighbors by self-coordination.

Light-path assessment is to determine the value of $\Omega$ given local information $X_i$, $Y_i$ at the first domain and average information $V_i$’s at the other domains. Figure 4 shows the estimation of the average domain-centric information as the first stage, where links illustrate information exchange among domains. The average information can be regarded as a function $f_1()$ acting on $X_i$ as well as information from other domains for $1 \leq i \leq L$. The path assessment is shown as the second stage, where decisions are made based on local information from the first domain and the average information from other subnets.

![Figure 4. Information estimation and decision on light-path availability](image)

The questions of our interest are:

(a) What is the average domain-centric information $V_i$ for $1 \leq i \leq L$ which optimizes the decision on path assessment?

(b) How to obtain $V_i$ by exchanging information among domains? How many domains need to be involved in the information-exchange? What is the information exchanged?

(c) What is the performance of light-path assessment using the local and average information compared to that of the optimal?

IV. OPTIMAL ASSESSMENT

We begin with the optimal assessment which provides a basis on what information needed to achieve the best assessment.

A. Optimal Assessment

Given local observations $X_i$ and $Y_i$, the light-path assessment can be cast into a decision problem [24]. The optimal decision results from the Bayes rule which determines

\[ \Omega = 1, \text{ if } P(I = 1|X_i,Y_i) \geq 1 - P(I = 1|X_i,Y_i), \]

\[ \Omega = 0, \text{ otherwise}, \]

where $I = (I_1,...,I_L)$. Let $I_{\text{max}}$ be the maximum number of intra-domain calls among all links within this domain. The Bayes rule can be rewritten as

\[ \Omega = 1, \text{ if } u(F_i - X_i - Y_{\text{max}})P(I' = 1|X_i) > \frac{1}{2}, \]

\[ \Omega = 0, \text{ otherwise}, \]

where $I' = \prod_{i=2}^{L} I_i$. $u(x)$ is an indicator function with $u(x) = 1$, if $x > 0$, and $u(x) = 0$, otherwise. $u(F_i - X_i - Y_{\text{max}})$ characterizes the availability of wavelengths at the head domain, where a meaningful case to consider is $F_i - X_i - Y_{\text{max}} > 0$, corresponding to the local availability of a wavelength.

$P(I' = 1|X_i)$ denotes the availability of the path downstream given local observations. Note that the conditional probability is now expressed only on $X_i$ as inter-domain calls introduced dependence among domains. $P(I' = 1|X_i)$ denotes the global information, and can be obtained by probing the path downstream. Such a scheme performs global modeling of a path, and would be suitable if all domains vary randomly and nearly simultaneously. If inter- and intra-domain calls vary more at a certain domains and less at the others, it would be more natural to consider domain-centric average information which decomposes the joint probability into independent factors relating to individual $I_i$’s.

The feasibility of such an approximation depends on inter-domain traffic that results in statistical dependence among subnets. For example, if there is no inter-domain traffic and intra-domain traffic is statistically independent across domains, the factorization would be exact, i.e., $P(I_2 = 1,...,I_L = 1|X_i) = P(I_2 = 1)...P(I_L = 1)$. If wavelength occupancy for inter-domain calls is statistically dependent across domains, the accuracy of such a factorization would depend on what the traffic pattern is and how the dependence among domains is taken into consideration. Therefore, a bottom-up approach is needed to relate inter-domain traffic at the WDM layer with domain-centric information estimations and exchange at management layer.

B. Traffic Model

We begin with a traffic model that is feasible for factorization. Such a model was developed in [30], and extended to inter-domain traffic in our prior work [1]. For simplicity, the model makes the following assumptions [1].

(i) The inter-domain calls originate and exit only at edge wavelength converters.

(ii) An inter-domain call exits subnet $i$ with probability $P_i$, and continues to the next subnet with probability $1 - P_i$.
(iii) A new inter-domain call enters subnet \(i\) with probability \(P_{in}\) on a free wavelength.

(iv) Wavelengths used by intra-domains calls are independent across different wavelengths and different links within a domain (and certainly across domains).

(vi) In domain \(i\), a wavelength is used for a local call in a link with probability \(\rho_{ii}\), and for an inter-domain call with probability \(\rho_{ii}'\). The probability that a wavelength is used for either a local or an inter-domain call is \(\rho = \rho_{ii} + \rho_{ii}'\). We define \(\alpha = \rho_{ii}'/\rho\), which characterizes the percentage of wavelengths occupied by inter-domain calls in domain \(i\).

Assuming the usage of wavelengths of different nominal values are independent, it can be shown similar to [1] that the number of inter-domain calls collected at domain boundaries can be approximated in distribution as joint Gaussian random variables. The distribution has a mean vector \(\mu = [\mu_1, ..., \mu_L]\), and a covariance matrix \(\Sigma\), where

\[
\mu = (F_{i} \rho_{2i}, \rho_{3i})
\]

\[
\Sigma^{-1} = \begin{bmatrix}
    d_{11} & d_{12} & 0 & \cdots & 0 & 0 \\
    d_{12} & d_{22} & d_{23} & \cdots & 0 & 0 \\
    0 & d_{32} & d_{33} & \cdots & \vdots & \vdots \\
    \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & 0 & \cdots & d_{L-1, L-1} & d_{L, L} \\
    0 & 0 & 0 & \cdots & 1 & 0
\end{bmatrix}
\]

\[
d_{ii} = \begin{cases} 
    \frac{1}{(1-\rho_{ii}^2)\sigma_i^2}, & \text{for } i = 1, \\
    \frac{1}{(1-\rho_{i+1,i}^2)\sigma_i^2}, & \text{for } i = L, \\
    \frac{1}{(1-\rho_{i+1,i}^2)\sigma_i^2} + \frac{\rho_{i+1,i}^2}{(1-\rho_{i+1,i}^2)\sigma_i^2}, & \text{otherwise}; \\
\end{cases}
\]

\[
d_{ij} = -\frac{\rho_{ij}}{\sigma_i \sigma_j}, \quad \rho_{ij} \text{ is the correlation coefficient between } X_i \text{ and } X_j \text{ for } i \neq j, 1 \leq i, j \leq L.
\]

C. A Simple Probabilistic Graphical Model

The tri-diagonal inverse-covariance matrix shows that the inter-domain calls form a (one-dimensional) Gaussian random field [26] whose statistical dependence can be represented by a simple probabilistic graph shown in Figure 5. Node \(X_i\) in the graph represents the random number of inter-domain calls at subnet \(i\), and the links show the conditional independence of \(X_i\) given its neighbors. For example, \(X_i\) is linked with \(X_{i-1}\) and \(X_{i+1}\), showing that \(X_i\) is conditionally independent of all the other inter-domain calls given those of its nearest neighbors, i.e., \(P(X_i | X_j, \forall j \neq i) = P(X_i | X_{i-1}, X_{i+1})\). Hence, \(X_i\) for \(1 \leq i \leq L\) form a special case (i.e., one-dimension) of Markov Random Field (MRF) [2][4]. \(X_i\) together with the decision variable \(I_i\) for \(1 \leq i \leq L\) result in a Hidden Markov model [25].

![Figure 5. Probabilistic graph for inter-domain calls and the availability of wavelengths at domains.](image)

The conditional independence of the random variables allows a decomposition of the connected graph into an unconnected graph shown in Figure 6. Here \(N_i\) for \(1 \leq i \leq L\) are independent random variables but take the statistical dependence among \(X_j\)’s \((j \neq i)\) into consideration.

![Figure 6. Decomposition of the Hidden Markov Model.](image)

D. Optimal Mean-Field Approximation

In this work, such a decomposition is equivalent to factoring \(P(I_2 = 1, ..., I_L = 1 | X_1)\) as

\[
P(I_2 = 1, ..., I_L = 1 | X_1) = \prod_{i=2}^{L} Q_i(X_i) + \epsilon, \quad (9)
\]

where \(Q_i\) characterizes the availability of a wavelength at domain \(i\) and \(\epsilon\) is the error due to the factorization. The optimal distributed estimation is to obtain \(Q_i(X_i)\) for \(1 \leq i \leq L\) so that the error \(\epsilon\) is minimized.

When there are no inter-domain calls, \(\epsilon = 0\), and

\[
P(I_2 = 1, ..., I_L = 1 | X_1) = \prod_{i=2}^{L} P(I_i = 1), \quad (10)
\]
where $Q_i(X_i) = P(I_i = 1)$ for all $i$ are the actual marginal probability of wavelength availability in domain $i$.

When there are inter-domain calls which enter and exit different subnets, $Q_i(X_i)$’s should not be the marginal probabilities $P(I_i = 1)$’s but should take into account the dependence among $X_j$’s to approximate the joint probability accurately. This would result in a set of Mean-Field equations [2]. Those equations are nonlinear and coupled but can be solved iteratively.

These Mean-Field equations, however, depend on the number of inter-domain calls $X_i$ at the head domain. This means that the optimal decomposition varies with respect to $X_i$, and thus requires solving Mean-Field equations for its given value. This is computationally infeasible especially when the number of wavelengths is large. Hence, a sub-optimal scheme for the decomposition is needed.

V. DISTRIBUTED INFORMATION ESTIMATION

A sub-optimal scheme performs distributed information-estimation by optimizing a criterion which depends on the average rather than the instantaneous wavelength usage at domains. This involves solving only one set of Mean-Field equations.

A. Mean-Field Approximation

The sub-optimal Mean-Field approximation is equivalent to decomposing the non-blocking probability of the path into independent factors,

$$P(I_1 = 1, ..., I_L = 1) = \prod_{i=1}^{L} Q_i + \epsilon',$$

where $Q_i$ for $1 \leq i \leq L$ correspond to non-blocking probabilities of individual domains which do not depend on values of $X_i$, and $\epsilon'$ is the approximation error. The Mean-Field approximation then involves deriving a set of equations for $Q_i$’s which would minimize $\epsilon'$.

A commonly-used criterion for deriving $Q_i$’s is the Helmholtz free energy or K-L information divergence [2][27]. Such a criterion minimizes the error due to the factorization in an information theoretical sense, and serves as a bound for the approximation error. The bound, however, becomes loose when inter-domain calls are strongly correlated, resulting in an inaccurate approximation. The reason for this is due to the normalization constants in the criterion which increase rapidly with the correlation among inter-domain calls. This suggests that a more accurate criterion than the Helmholtz energy is needed to obtain $Q_i$’s.

Such a criterion is motivated by a Mean-Field expansion of $P(I = 1)$ [27]. The expansion was initially developed for obtaining a higher-order Mean-Field approximation [27]. The expansion, however, can not be used directly in our case, due to the normalization constants just described. Hence we extend the original Mean-Field approach to the Fourier domain.

Our starting point is similar to that used in [27][28]. But the joint Gaussian distribution of inter-domain calls $N(\mu, \Sigma)$ is represented through the inverse Fourier transform of its characteristic function. In particular, we have

$$P(I = 1) = \ln \int \prod_{i=1}^{L} (1 - a_i e^{-\rho_i - M_i}) \frac{1}{2\pi} e^{-\frac{1}{2} \sum_{j=1}^{L} \omega_j^2} d\omega dM,$$

where $e^{-\frac{1}{2} \sum_{j=1}^{L} \omega_j^2}$ is the characteristic function of a density function $N(0, \Sigma)$, $M_i = (M_1, ..., M_L)$ is a vector of integration variables, and $\omega_j = (\omega_1, ..., \omega_L)$ represents the transpose of vector $\omega$ in the complex plane.

$P(I = 1 | M_i) = \prod_{j=1}^{L} (1 - a_i e^{-\rho_i - M_i})$ was derived in our prior work for centralized assessment, where $a_i = 1 - (1 - \rho_i)^N$ is the probability that a wavelength is used in domain $i$ [1], and $1 - a_i e^{-\rho_i - M_i}$ is the non-blocking probability of domain $i$ given the number of inter-domain calls $M_i$.

Observing that the dependence among $X_i$’s results from the off-diagonal elements in $\Sigma$, we introduce a function of a variable $\beta$ ($0 \leq \beta \leq 1$),

$$H(\beta) = \ln \left( \prod_{i=1}^{L} (1 - a_i e^{-\rho_i - M_i}) \right)$$

$$\frac{1}{2\pi} \int e^{-\frac{1}{2} \sum_{j=1}^{L} \omega_j^2} \frac{1}{2} (z & - \Lambda) = \int e^{-\frac{1}{2} \sum_{j=1}^{L} \omega_j^2} (w \cdot \sigma)^T (z & - \Lambda) = \int e^{-\frac{1}{2} \sum_{j=1}^{L} \omega_j^2} \omega^T (w \cdot \sigma) \right) d\omega dM,$$

where $\Lambda$ is a diagonal matrix with the same diagonal elements of $\Sigma$, $\gamma(\beta) = \{ \gamma_1(\beta), ..., \gamma_L(\beta) \}$ is an unknown vector depending on $\beta$, and $\sigma$ is a statistical average of $\omega$ (see [27] for details).

To understand the roles of $\beta$ and $\gamma(\beta)$, we consider two special cases. When $\beta = 1$, $\gamma(1) = 0$, $H(1) = \ln P(I = 1)$, which is the logarithm of the original probability. When $\beta = 0$, $H(0) = \ln \prod_{i=1}^{L} P(I_i = 1)$, which corresponds to the case that the wavelength usage at different domains is independent. Observe that in the characteristic function, the term $\frac{1}{2} (z & - \Lambda) \omega$ characterizes the dependence among the random variables, whereas $\frac{1}{2} \omega^T (w \cdot \sigma)$ contributes to a non-zero mean. Therefore, a proper choice of $\gamma(\beta)$ should minimize the error when $\frac{1}{2} \omega^T (w \cdot \sigma)$ is used to account for the statistical dependence of $X_i$’s in replacement of $\frac{1}{2} (z & - \Lambda) \omega$. In other words, the spatial dependence of inter-domain calls can be absorbed in an adjusted mean vector depending on $\gamma(\beta)$. As the result, $P(I = 1)$ can be approximated by a product of independent factors.
We choose dominant terms in the Taylor expansion of 
\[ \ln(P(I=1) = \beta = 0 \] [27]. For simplicity, the first two terms in the Taylor expansion are chosen, where
\[ H_i(1) = H(0) + H(0). \] (12)

\( H(0) \) corresponds to the decomposition of the conditional probability \( P(I=1) \) by ignoring the dependence among domains, and \( H(0) \) takes the dependence into consideration.

**Lemma 1:**
\[ H_i(1) = \sum_{i=1}^{L} \ln v_i - \sum_{i=1}^{L} \chi' \left( \frac{1}{v_i} - 1 \right) \ln a_i \]
\[ + \frac{1}{2} \sum_{i=1}^{L} \sum_{k=L+1}^{L} c_{ik} \left( \frac{1}{v_i} - 1 \right) \ln a_i \ln a_k, \] (14)

where \( v_i = 1 - a_i e^{-\beta v_i} \), for \( 1 \leq i \leq L \).

The proof of Lemma 1 is given in Appendix 1, which involves calculations of the derivatives in the Taylor expansion. The Mean-Field approximation aims to find \( \gamma(\beta) \) by maximizing \( H_i(1) \). This results in a set of coupled Mean-Field equations.

**Theorem 1:**
The Mean-Field equations can be obtained by maximizing \( H_i(1) \) with respect to \( \gamma \), for \( 1 \leq i \leq L \):
\[ \gamma_i = -\frac{1}{d_{ii}} \sum_{k=1}^{L} d_{ik} \left( \gamma_k + c_{ik} \frac{1 - v_i}{v_k} \ln a_k \right) \]
\[ + \frac{(1 - d_{ii}) c_{ii} - 1 - v_i}{v_i} \ln a_i, \] (15)

where \( d_{ik} \) is the element at the \( i \)th row and the \( k \)th column of \( \Sigma^{-1} \) for \( 1 \leq i, k \leq L \).

The proof of the theorem is given in Appendix 2, and obtained by setting \( \partial H_i(1) / \partial \gamma = 0 \). As \( v_i \) depends on \( \gamma \), the Mean-Field equations are nonlinear and coupled but can be solved iteratively. The coupling among \( \gamma \) results from the spatial dependence introduced by the inter-domain calls and exhibited in the graphical model.

**Theorem 2:** The factorized probability which maximizes \( H_i(1) \) is
\[ \Pi_{i=1}^{L} Q_i = \Pi_{i=1}^{L} \frac{\gamma_i}{v_i a_i} e^{-\beta v_i} \] (16)

where
\[ Q_i = v_i a_i e^{-\beta v_i}. \]

The proof of the theorem is given in Appendix 3, and obtained through combining Equations (14) and (15). \( \Pi_{i=1}^{L} Q_i \) is the Mean-Field approximation of the non-blocking probability \( P(I=1) \). Each \( Q_i \) can be regarded as an approximation to the non-blocking probability of “independent” domain \( i \), but accounts for the dependence among random variables \( X_1,...,X_L \) through \( \gamma \), for \( 1 \leq i \leq L \). \( \gamma \) is an adjustment to the mean value of \( X_i \) so that the decomposition can take the statistical dependence into consideration.

**B. Distributed Algorithm**
A distributed algorithm finds iteratively a solution for the Mean-Field equations. Let \( n_i \) be a neighborhood of \( i \), i.e., \( n_i = \{i-1,i+1\} \) for \( 2 < i < L \), \( n_i = \{i+1\} \), for \( i = 1 \), and \( n_i = \{i-1\} \), for \( i = L \).

**Algorithm:**
1. Initialize \( i = 1 \), choose \( \delta > 0 \) as a stopping threshold.
2. The \( i \)th domain calculates the second-order statistics \( d_{ii} \) and \( d_{ik} \) for \( k \in n_i \), which is assumed to be static during the Mean-Field updates.
3. The \( i \)th domain obtains average information \( \gamma_i(t), v_i(t) \) from its neighboring domains \( k \in n_i \) at time \( t \), and provides updates \( \gamma_i(t+1), v_i(t+1) \) to neighbors.
   4. \( i = i + 1 \).
   If \( i < L \), go to step 1;
   Else if \( i = L \), let \( \theta(t+1) = |\gamma_i(t) - \gamma_i(t+1)| \).
      If \( \theta(t+1) \leq \delta \), go to (5);
      Otherwise, \( i = 1 \), go to (1).
5. Stop.

**C. Decision Rule on Path Availability**
Using the estimated information, the path availability can be determined through a simple decision rule as follows.
1. Domain \( i \) sends \( Q_i \) to Domain \( j \), \( 2 \leq i \leq L \). Domain \( j \) computes \( \Pi_{i=2}^{L} Q_i \) as an approximation to the joint probability \( P(I_2 = 1,...,I_L = 1) \), of the path availability downstream.
2. Domain \( L \) also has \( \gamma_L \) as the adjusted mean and \( \sigma^2_L \) as the variance of the inter-domain calls within its own domain.
\[ \Omega = 1, \text{ if } u(F_i - X_i - Y_{\text{max}}) a_{\text{k1}} \frac{N(\mu_1 + \gamma_1, \sigma^2_1)}{N(\mu, \sigma^2)} \prod_{i=2}^l Q_i > \frac{1}{2}, \]

\[ \Omega = 0, \text{ otherwise}, \]

where \( N(\mu, \sigma^2) \) is the normal distribution with mean \( \mu \) and variance \( \sigma^2 \).

This decision rule can be explained as follows: The expected value (see the beginning of Appendix 4 for details)
\[
E\left[u(F_i - X_i - Y_{\text{max}}) a_{\text{k1}} \frac{N(\mu_1 + \gamma_1, \sigma^2_1)}{N(\mu, \sigma^2)} \prod_{i=2}^l Q_i \right] = \prod_{i=2}^l Q_i.
\]

That is, \( u(F_i - X_i - Y_{\text{max}}) a_{\text{k1}} \frac{N(\mu_1 + \gamma_1, \sigma^2_1)}{N(\mu, \sigma^2)} \prod_{i=2}^l Q_i \) can be considered as an approximation of \( P(I = 1 | X_i) \). \( \prod_{i=2}^l Q_i \) characterizes the average wavelength usage at the path downstream. If \( u(F_i - X_i - Y_{\text{max}}) = 1 \), there is a wavelength available locally at domain \( I \). Then an additional condition needs to be included: for \( \gamma_i \neq 0 \) (in fact \( \gamma_i \leq 0 \) for \( c_u > 0 \), see Appendix 2 for details), the term
\[
\frac{N(\mu_1 + \gamma_1, \sigma^2_1)}{N(\mu, \sigma^2)} e^{-\frac{v_i}{\sigma^2}}
\]
favors a value of \( X_i \) satisfying
\[
0 \leq X_i \leq \mu_1 + \lambda. \quad \text{Meanwhile, } 0 < a_{\text{k1}} \frac{N(\mu_1 + \gamma_1, \sigma^2_1)}{N(\mu, \sigma^2)} e^{-\frac{v_i}{\sigma^2}} \leq 1 \text{ for } \gamma_i \leq 0.
\]

\[
a_{\text{k1}} \frac{N(\mu_1 + \gamma_1, \sigma^2_1)}{N(\mu, \sigma^2)} e^{-\frac{v_i}{\sigma^2}}
\]
thus shows how a decision based on the use of local wavelength usage is affected by the other domains downstream. That is, the local decision should be more conservative when the average impact from other domains is taken into consideration.

Note that when there are no inter-domain calls, \( \gamma_i = 0 \), \( \prod_{i=2}^l Q_i = P(I = 1) \), and \( u(F_i - X_i - Y_{\text{max}}) \prod_{i=2}^l Q_i > \frac{1}{2} \) is used in the decision.

As soon to be shown, such a simple decision rule can result in a reasonably accurate assessment if the average usage at domains is representative of wavelength occupancy.

VI. INFORMATION EXCHANGE

What is the information exchanged among domains required by the suboptimal distributed scheme? How does the amount of information exchange compare to that for the centralized and decentralized schemes? We now answer these questions by examining the information exchange in estimation and decision.

A. Distributed Information Estimation

Consider first the information exchange for estimation using the suboptimal distributed algorithm in Section V. B. The estimation process obtains \( Q_i \) that depends on \( \gamma_i \) from the coupled Mean-Field equations (15). The structure of the probabilistic graph shows explicitly which and how many neighboring domains should exchange information. For example, links among nodes \( i, i-1 \) and \( i+1 \) on the graph show that domain \( i \) needs to exchange information with its nearest neighbors during estimation. This explains why a bottom-up approach we choose is important since the inter-domain traffic determines the spatial dependence and thus the graphical structure.

The specific information exchanged results from the Mean-Field equations. In particular, during the estimation process, domain \( k \) (1 \( \leq k \leq L \)) sends to its neighbors the following parameters pertaining to this domain: (a) the information on wavelength usage, \( v_i \) and \( \log a_i \), (b) a correction to the mean \( \gamma_i \), and (c) the second-order statistics \( c_{ik} \) and \( d_{ik} \) of inter-domain traffic.

\( v_i \) and \( \log a_i \) provide the aggregated information on intra-domain traffic at domain \( k \). Recall \( v_i = 1 - a_i^{(F_i - Y_{\text{max}})} e^{-\frac{2}{\sigma^2}} \).

For \( \gamma_i = 0 \), \( v_i = 1 - a_i^{(F_i - Y_{\text{max}})} e^{-\frac{2}{\sigma^2}} \) is the marginal probability \( P(I = 1) \) (see Equation (20)). Intuitively, since a domain does not see the detailed fluctuation of intra-domain calls within other domains, such information would be needed for lightpath assessment. Meanwhile, as domain \( k \) is viewed as a node in the factorization, only average information at the domain-level, e.g. \( \log a_i \), needs to be exchanged with other subnets. This eliminates the need of exposing the detailed topology- and load-information within a domain.

\( \gamma_i \) is an adjustment to the average wavelength usage. The role of \( \gamma_i \) is to account for the spatial dependence when the joint probability is decomposed into independent factors. A non-zero value for \( \gamma_i \) means that the dependence among wavelength usage of inter-domain calls is taken into consideration in an average sense. \( \gamma_i \) is updated in solving the Mean-Field equations. The update can be asynchronous, suggesting that the average information can be sent based on local fluctuations at a domain. The update only requires information from neighboring domains. This results from the conditionally-independent nature of inter-domain calls, where the nearest neighbors of a domain contain sufficient information about the farther neighbors. As explained earlier, such an information exchange is shown clearly by the structure of the graphical model.

\( c_{ik} \) and \( d_{ik} \) characterize the second-order statistics of inter-domain calls. They can be evaluated either locally or through exchanging information with neighboring domains. These quantities are assumed to be static when the Mean-Field equations are being solved.

In summary, the information is exchanged among neighbors during estimation. In particular, the information corresponds to a set of parameters \( V_i = \{ v_i, \log a_i, c_{ik}, d_{ik}, \gamma_i \} \),
where \( \{v_i, \gamma_i\} \) is dynamic and updated during the estimation process, and \( \{\log a_i, c_i, d_i\} \) is static.

The information-exchange is global for deciding the path availability, where all subnets send \( Q_i (2 \leq i \leq L) \) to domain \( I \). Such information needs to be transmitted once the Mean-Field solutions are obtained. The information can be further updated whenever local traffic fluctuations trigger an update.

B. Centralized and Decentralized Schemes

Two centralized schemes can be considered. One is for a decision at the head domain where global information on the path downstream is needed to estimate the conditional probability \( P(I' = 1 | X_i) \) for all values of \( X_i \) (Fig. 4). This can be implemented through probing, and the number of probes needed corresponds to the amount of information required. No additional information needs to be sent from domains to the head node during the decision process.

The other centralized scheme is similar to that discussed in the prior work [1] where \( X_i \)'s are collected at each subnet and sent to the head domain directly for decision making. \( X_i \) is a random variable and can take on any value in \( \{1,..., F_i\} \).

In the decentralized estimation, domain \( i \) obtains \( Q_i = P(I_i = 1) \) which is the exact marginal non-blocking probability. Such information is available within a domain and thus no information-exchange takes place during estimation. \( Q_i \)'s are then sent to the head domain for path assessment. The decision rule is a special case of the distributed decision with \( \gamma_i = 0 \):

\[ \Omega = 1, \text{ if } u(F_i - X_i - Y_{\max}) \prod_{i=2}^{L} Q_i > \frac{1}{2}, \]

\[ \Omega = 0, \text{ otherwise}. \]

Combining information-exchange in both estimation and decision, the centralized scheme requires the most information exchange, and the decentralized scheme requires the least, while the distributed scheme is between these two extremes.

VII. PERFORMANCE

As any sub-optimal algorithm can be chosen based on heuristics, it is important to study a distributed algorithm whose performance is predictable and within a tolerable degradation from that of the optimal scheme [15]. In this section, we study the performance of the distributed estimation and compare that with the centralized and decentralized schemes.

A. Centralized Scheme

For the centralized scheme which uses \( X_i \) (1 \( \leq i \leq L \)) directly as the domain-centric information [1], the probability \( P_{se} \) of an incorrect assessment satisfies

\[ P_{se} \leq P_{cr} \leq \min\{P_b, 1 - P_b\}, \quad \text{(17)} \]

where \( P_b \) is the blocking probability of the path.

Using a similar method to that in [1], we can show that the same inequality holds for the optimal assessment in Equation (2).

\[ P_{se} \leq P_{cr} \leq \min\{P_b, 1 - P_b\} + 3\Delta. \quad \text{(18)} \]

The proof is given in Appendix 4, and obtained by comparing with the optimal centralized assessment. The theorem shows that the upper bound differs from that of the optimal scheme by a small quantity proportional to \( \Delta \). The approximation error \( \Delta \) is introduced by the factorization of the Mean-Field decomposition. When all \( I_i \), 1 \( \leq i \leq L \), are independent, \( \Delta \) is zero and the equalities hold for both the lower and the upper bounds. When \( I_i \), 1 \( \leq i \leq L \), are dependent, \( \Delta \) should vary with respect to the dependence introduced by the inter-domain traffic.

B. Distributed Scheme

Assume decisions are now made on path assessment using the average information estimated from the distributed algorithm.

**Theorem 3:**

Let \( P \) be the probability of an incorrect assessment using the suboptimal scheme given in Section V. C. Let \( \Delta = \text{Pr}(I = 1) - \prod_{i=2}^{L} Q_i \) be the approximation error which is the difference between the non-blocking probability and its Mean-Field approximation. Then

\[ P_{se} \leq P \leq \min\{P_b, 1 - P_b\} + 3\Delta. \quad \text{(19)} \]

The proof is given in Appendix 4, and obtained by comparing with the optimal centralized assessment. The theorem shows that the upper bound differs from that of the optimal scheme by a small quantity proportional to \( \Delta \). The approximation error \( \Delta \) is introduced by the factorization of the Mean-Field decomposition. When all \( I_i \), 1 \( \leq i \leq L \), are independent, \( \Delta \) is zero and the equalities hold for both the lower and the upper bounds. When \( I_i \), 1 \( \leq i \leq L \), are dependent, \( \Delta \) should vary with respect to the dependence introduced by the inter-domain traffic.

C. Decentralized Scheme

Decentralized estimation is a special case of distributed schemes where no information-exchange among domains during the estimation process. This corresponds to the situation that individual domains choose to ignore completely the dependence of inter-domain calls, and simply use the marginal probabilities in the factorization. That is, \( Q_i = P(I_i = 1) \), is the average domain information for 1 \( \leq i \leq L \). The specific form of the marginal probability can be evaluated for 1 \( \leq i \leq L \), where

\[ Q_i = P(I_i = 1) = 1 - a_i^{(E_i - \mu)} e^{\frac{\mu_i^2 a_i^2}{2}}, \quad \text{(20)} \]

and

\[ \prod_{i=2}^{L} P(I_i = 1) = \prod_{i=2}^{L} (1 - a_i^{(E_i - \mu)} e^{\frac{\mu_i^2 a_i^2}{2}}). \quad \text{(21)} \]
D. Performance Comparison

We now compare the performance of path assessment due to the centralized, distributed and decentralized schemes for information-estimation. Ideally, such comparisons should be made on the probability of assessment error. But the actual probability of incorrect assessment of the centralized scheme has been shown to be close to the upper bound (see [1] for details). Therefore, we compare the upper bounds in terms of the exact blocking probability for the centralized scheme, the Mean-Field approximation for the distributed estimation, and the product of marginal probabilities for the decentralized scheme.

To understand how the differences may vary with respect to the spatial dependence of inter-domain traffic, we first examine the relationship between the true probability of path availability \( P(I = 1) \) and the correlation coefficients of \( X_i \)'s. The proposition below shows that \( P(I = 1) \) increases with the correlation coefficients.

**Proposition 1:**

\[
P(I = 1) \text{ is an increasing function of covariance } c_{ik} \text{ for } i \neq k, 2 \leq i \leq L. \text{ In particular, the rate of increase can be characterized as}
\]

\[
\frac{\partial P(I = 1)}{\partial c_{ik}} = \ln a_i \ln a_k
\]

\[
= \left( a_i^{(E_j - \mu_j - M_j)} a_k^{(E_k - \mu_k - M_k)} \prod_{j \neq i, k} (1 - a_j^{(E_j - \mu_j - M_j)}) \right) N(0, \Sigma) dM
\]

which is positive for \( c_{ik} \neq 0 \).

The proof can be obtained through direct calculations of the straightforward derivative of \( P(I = 1) \), and is thus omitted. This suggests that the factorization would deviate more and more from the true path availability when the correlations increase among inter-domain calls. Now consider the distributed scheme due to the Mean-Field approximation.

**Proposition 2**

Assume the covariance \( c_{ik} \geq 0, i \neq k, 1 \leq i, k \leq L \). Then each \( Q_i \) in the mean-field solution (15) satisfies,

\[
Q_i > P(I = 1) \text{ for } c_{ik} \geq 0; \quad i \neq k, 1 \leq i, k \leq L; \text{ and}
\]

\[
Q_i = P(I = 1) \text{ for } c_{ik} = 0; \quad i \neq k, 1 \leq i, k \leq L;
\]

for all \( 1 \leq i \leq L \), where the \( P(I = 1) \) is the marginal probability.

The proof of the proposition is given in Appendix 4, and can be obtained by showing \( \gamma_i < 0 \) for \( 1 \leq i \leq L \) based on algebraic manipulations. This result suggests that the Mean-Field solution and the exact blocking probability may have the same trend when the dependence among domains increases. But the marginal probabilities from the decentralized scheme do not vary with the correlation.

We now compare the performance of the three schemes numerically. Let \( P_a \) be the probability that a new inter-domain call enters subnet \( i \). An existing inter-domain call continues to the \( ith \) domain with probability \( P_i \) from the previous domain. When \( P_a = P_i = P_i \) for all \( 1 \leq i \leq L-1 \). Then we have

\[
\rho_{i,i+1} = 1 - \frac{P_i}{P_2}, \quad \forall 1 \leq i \leq L-1.
\]

Figures 7 and 8 plot the correlation coefficient \( \rho _g \) versus the estimated path-availability probabilities for the three schemes. Assume uniform network domains. \( F \) is the number of wavelength; \( L \) is the number of domains; \( \rho \) is the load of total wavelength load; \( H \) is the number of hops within a domain; and \( \alpha \) is the probability that an occupied wavelength is used for inter-domain calls. For example, in Figure 7, there are 3 network domains, each of which has 8 links and each link supporting 40 wavelengths. A wavelength is used in each network with probability 0.8, and 90% of the occupied wavelengths are used for inter-domain calls. It can be observed that when \( \rho _g \) increases, i.e., the dependency among network domains becomes stronger, the decentralized scheme underestimates the path availability probability more and more significantly. When \( \rho _g \) is close to one, the decentralized scheme introduces approximately 10% of error. On the other hand, the distributed estimation is able to follow the true path blocking probability accurately.

Note that the improvement can be pronounced using distributed estimation for high wavelength load where the path-availability probability is small. This shows that the information exchanged seems to be more informative when the path is heavily loaded.
that of the optimal due to the accuracy in the estimation process as shown in Figures 7 and 8.

![Figure 8. Estimation of the probability of path-availability; F=100, L=12, \( \rho = 0.8, \alpha = 0.9, H = 8 \).](image)

Figures 9 and 10 depict two examples of the upper bound of the probability error for the optimal centralized scheme and the suboptimal distributed scheme. It is observed that the distributed scheme is able to achieve a performance close to that of the optimal due to the accuracy in the estimation process as shown in Figures 7 and 8.

![Figure 9. Comparison of the upper bound for the probability of error; F=40, L=3, \( \rho = 0.8, \alpha = 0.9, H = 8 \).](image)

![Figure 10. Comparison of the upper bound for the probability of error; F=100, L=12, \( \rho = 0.8, \alpha = 0.9, H = 8 \).](image)

The computational cost is an important issue to the practicality of the distributed estimation. The cost depends on the number of iterations needed in solving the Mean-Field equations. Figure 11 depicts the number of iterations as a function of the correlation coefficient among inter-domain wavelength occupancy. It is observed that the number of iterations increases with the correlation coefficient \( \rho_x \). This is because the Mean-Field approach tries to decompose the path blocking probability into independent factors associated with domains. This decomposition becomes more and more difficult when the spatial dependence among wavelength occupancy at domains becomes stronger. When \( \rho_x \) falls into a reasonably range, i.e., \( 0 < \rho_x < 0.95 \), the number of iterations is bounded by 60. When \( \rho_x \) is close to 1, the number of iterations can exceed 100. Note that the convergence time may be reduced when more sophisticated iterative algorithms are used.

**VIII. CONCLUSIONS**

In this work, we have studied network management information for distributed assessment of light-path availability across administrative domains. Our investigation is based on both applications of graphical models and the Mean-Field theory in machine learning. Using these approaches, we have obtained the following findings when the aggregated inter-domain traffic exhibits spatial Markov dependence.

(a) A global measure of end-end path availability can be approximated using instantaneous wavelength usage at the head domain and average domain-centric information from other subnets.

(b) The average domain-centric information can be estimated through exchanging several parameters among neighboring domains. The number of neighbors involved in the information-exchange is determined by the structure of the probabilistic graphical model, which is either 2 or 1 in this work. The actual average information exchanged during estimation is determined by the distributions of inter-domain and intra-domain calls, topology and resource of the domains.
In this work, Gaussian approximation is used for inter-domain traffic. Yet other traffic patterns with the same spatial dependence would also allow the use of the graphical structure.

Domains can update the average information asynchronously and iteratively based on a distributed algorithm which solves the Mean-Field equations.

(c) A simple decision rule can then be used by the head domain to determine the path availability for an end-end request. The performance degradation of such a decision rule from that of the optimal can be quantified by an error due to the Mean-Field approximation.

(d) The information-exchange in the estimation stage seems to be pronounced for improving the performance under heavy network-load conditions. The current work is limited to one path, one traffic pattern, and static traffic. Extensions of this work include multiple paths where the impact of network topologies and other traffic patterns can be studied. Meanwhile, dynamic inter-domain traffic should be interesting to explore along with asynchronous updates of the parameters. Although the current work is for optical networks, the approaches may also be applied to other network settings.

ACKNOWLEDGMENT

The authors would like to thank Sung-eok Jeon, Zesheng Chen, Anwar Elwalid, and Phil Whiting for helpful discussions. Support from NSF ECS 0300605 (ECS 9908578) is gratefully acknowledged.

Appendix 1: Proof of Lemma 1

Proof:

By definition,

\[ H(0) = \ln \left( \prod_{d=1}^{D} (1 - a_i^{\mu_{-\mu_d} - M}) \right) + \frac{1}{2\pi} \int e^{-\frac{1}{2} \frac{(x,y)}{\sigma^2}} dxdy, \]

where \( \mu_i = E[X_i] \). Through definition of characteristic function and algebraic manipulation, we have

\[ H(0) = \sum_{i=1}^{L} \ln v_i - j \sum_{i=1}^{L} \gamma_i < \omega >, \]

where \( \gamma_i = \gamma_i(0) \) for simplicity, \( v_i = 1 - a_i^{\mu_i - \mu_d - M} \), \( \sigma^2 = \text{Var}(X_i) \) is the variance of \( X_i \), \( < \omega > \) is the conditional expectation of \( \omega \).

\[ < \omega > = \int \left( 1 - a_i^{\mu_i - \mu_d - M} \right) \frac{1}{2\pi} \int e^{-\frac{1}{2} \frac{(x,y)}{\sigma^2}} dxdy. \]

\[ j \int \left( 1 - a_i^{\mu_i - \mu_d - M} \right) \frac{1}{2\pi} \int e^{-\frac{1}{2} \frac{(x,y)}{\sigma^2}} dxdy = \frac{H(0)}{v_i}, \]

\[ = - j(\ln a_i) \frac{1 - v_i}{v_i}. \]

\( H(0) \) corresponds to the decomposition of the conditional probability \( P(I=1) \) by ignoring the dependence among domains. Now if the dependence is taken into consideration through the first-order term, we have

\[ H(1) = \sum_{i=1}^{L} \ln v_i - j \sum_{i=1}^{L} \gamma_i < \omega > + \frac{1}{2} \sum_{i=1}^{L} \sum_{k=i+1}^{L} c_{i,k} < \omega > < \omega >. \]

Inserting the expression of \( < \omega > \), we prove the theorem.

Q.E.D

Appendix: Proof of Theorem 1

Proof:

For \( 1 \leq i \leq L \), letting \( \frac{dH_i(1)}{d\gamma_i} = 0 \), we have the fix point equation for \( \lambda_i \),

\[ \frac{v_i}{v_i} + \ln a_i \frac{\ln a_i}{v_i} + \gamma_i \frac{\ln a_i}{v_i} + \sum_{k=i+1}^{L} c_{i,k} \ln a_i \frac{1 - \ln a_i}{v_i} \frac{-v_i}{v_i} = 0 \]

Inserting \( \gamma_i = \frac{1 - v_i}{v_i} \ln a_i \), we have

\[ \gamma_i + c_{i,k} \frac{1 - v_i}{v_i} \ln a_i = \sum_{k=i+1}^{L} c_{i,k} \frac{1 - v_i}{v_i} \ln a_i. \]

Now let \( G = (g_i)_{L \times 1} \), where \( g_i = \frac{1 - v_i}{v_i} \ln a_i \). The the above equation can be written in a matrix form,

\[ \gamma + \Lambda G = \Sigma G, \Rightarrow \Sigma^{-1} \lambda + \Sigma^{-1} \Lambda G = G, \]

\[ \Rightarrow \Sigma^{-1} \lambda = (\Sigma^{-1} \Lambda G + I) G. \]

This matrix form gives us coupled Mean Field Equations for \( 1 \leq i \leq L \).
\[ \gamma_i = -\frac{1}{d_{ik}} \sum_{k=1}^{n} d_{ik} \gamma_k + \frac{1}{d_{ii}} \sum_{i=1}^{L} (1-c_{ik} d_{ik}) \frac{1-v_i(\gamma_i)}{v_i(\gamma_i)} \ln a_i. \]

Note that the coupling is among the \( i \)-th, \((i-j)\)-th, and \((i+j)\)-th equations for \( 1 \leq i \leq L \), between \( i \)-th and \((i+j)\)-th for \( i=2 \), \( i \)th and \((i+j)\)-th for \( i=L-1 \). Note also that by rewriting the above expressions, we have

\[ \gamma_i = \sum_{k=1}^{n} c_{ik} \frac{1-v_i}{v_i} \ln a_i. \]

This shows that \( \gamma_i \leq 0 \), for \( c_{ik} > 0 \). This shows the sign of \( \gamma_i \) but is not feasible for its evaluation since the global information would be needed.

\[
\text{Appendix 3: Proof of Theorem 2}
\]

\[
v_i = 1-a_i^{2-e^{-\mu_i-K_i} e^x \frac{m_i^2 b_i^2}{a_i}},
\]

\[
H(1)_{\text{min}} = \sum_i \ln v_i - \sum_{i,k} c_{ik} \frac{1}{v_k} \left( \frac{1}{v_i} - 1 \right) \ln a_i \ln a_i
\]

\[ + \frac{1}{2} \sum_{i,k} c_{ik} \left( \frac{1}{v_k} - 1 \right) \left( \frac{1}{v_i} - 1 \right) \ln a_i \ln a_i
\]

\[ = \sum_i \ln v_i - \frac{1}{2} \sum_{i,k} c_{ik} \left( \frac{1}{v_k} - 1 \right) \ln a_i \ln a_i.\]

Now approximating \( \ln P(I=1) \) using \( H(1)_{\text{min}} \), we have

\[ Q_i = \nu_i e^{\frac{1}{2} \sum_{i,k} c_{ik} \left( \frac{1}{v_k} - 1 \right) \ln a_i \ln a_i} = \nu_i a_i^{\frac{1}{2} \frac{v_i}{v_k}}. \]

Hence,

\[ P(I=1) = \prod_{i=1}^L Q_i = \prod_{i=1}^L \nu_i a_i^{\frac{1}{2} \frac{v_i}{v_k}}. \]

\[ \text{Q.E.D} \]

\[
\text{Appendix 4: Proof of Proposition 2}
\]

\[ \text{Proof:} \]

For positive correlations of inter-domain calls, (which is the case here), \( \gamma_i < 0 \), \( 0 < \nu_i < 1 \). Then \( 0 < a_i^{\frac{v_i-1}{2} \frac{v_i}{v_k}} < 1. \)

\[ \frac{\partial \ln Q_i}{\partial \gamma_i} = \gamma_i \ln^2 a_i \left( 1 - \nu_i \right) < 0 \]

Since \( \gamma_i \leq 0 \), \( Q_i(\gamma_i) \geq Q_i(0) \), and \( Q_i(0) = P(I_i = 1) \). When \( c_{ij} \) increases, \( \lambda_i \) decreases, and \( Q_i(\lambda_i) \) increases. But \( P(I_i = 1) \) remains constant.

\[
\text{Appendix 5: Proof of Theorem 3}
\]

\[ \text{Proof:} \]

Let \( \hat{P}(I=11 X_i) \) be the approximated conditional probability for \( P(I=11 X_i) \) based on the simple decision rule in Section V.C. The probability of an incorrect assessment for a given \( X_i \) is

\[
P_e(X_i) = P(I=11 X_i) I(1-\hat{P}(I=11 X_i))
\]

\[ >P(I=11 X_i)) + (1-P(I=11 X_i)) I(1-\hat{P}(I=11 X_i)) \]

\[ \leq P(I=11 X_i)), \]

where \( I(A) \) is an indicator function, \( I(A)=1 \) if event \( A \) is true; and \( I(A)=0 \), otherwise. Then the probability of an incorrect assessment averaged over all \( X_i \) is

\[
P_e = E[P(I=11 X_i) I(1-\hat{P}(I=11 X_i))]
\]

\[ >P(I=11 X_i)) + (1-P(I=11 X_i)) I(1-\hat{P}(I=11 X_i)) \]

\[ \leq P(I=11 X_i)), \]

where the expectation is over \( X_i \). Replacing \( P(I=11 X_i) \) by

\[ \hat{P}(I=11 X_i) + |P(I=11 X_i) - \hat{P}(I=11 X_i)| \], and expand the expectation term by term, we have

\[
P_e = P_{el} + P_{eL},
\]

\[ P_{eL} = E[\hat{P}(I=11 X_i) I(1-\hat{P}(I=11 X_i))]
\]

\[ >P(I=11 X_i)) + E[(1-\hat{P}(I=11 X_i)) I(P(I=11 X_i))
\]

\[ >P(I=11 X_i))], \]

\[ P_{eL} = E[P(I=11 X_i) - \hat{P}(I=11 X_i)) I(1-P(I=11 X_i))
\]

\[ >P(I=11 X_i)) - E[(P(I=11 X_i) - \hat{P}(I=11 X_i)) I(\hat{P}(I=11 X_i))]
\]

\[ >P(I=11 X_i))]. \]

Now using a similar technique as Appendix 1 of [1], we have \( P_{el} \leq \min(E[\hat{P}(I=11 X_i))] - E[\hat{P}(I=11 X_i))] \).

Now \[ E[\hat{P}(I=11 X_i)] = \prod_{i=1}^L Q_i, \] and let

\[ \Delta = \prod_{i=1}^L Q_i - P(I=1) \] be the difference between the Mean-Field approximation and the true non-blocking probability. Then

\[ E[\hat{P}(I=11 X_i)] = P(I=1) + \Delta, \]

and

\[ \min(E[\hat{P}(I=11 M_i))] - E[\hat{P}(I=11 M_i))] \leq \min(1-P_{el}, P_{eL}) + |\Delta| \]

and therefore
\[ P_\varepsilon \leq \min(1 - P_b, P_g) + |\Delta| \cdot \]

Furthermore,
\[ E[P|I = 1|X_1] - \hat{P}(I = 1|X_1))|I - \hat{P}(I = 1|X_1)\]
\[ > \hat{P}(I = 1|X_1)) \leq |\Delta|, \]
\[ -E[|P(I = 1|X_1) - \hat{P}(I = 1|X_1)\hat{P}(I = 1|X_1)\]
\[ > 1 - \hat{P}(I = 1|X_1)) \leq |\Delta| \cdot \]

Hence,
\[ P_\varepsilon \leq 2|\Delta|. \]

Combining the upper bounds of \( P_\varepsilon \) and \( P_\varepsilon \) as an upper bound for \( P_\varepsilon \), we have
\[ P_\varepsilon \leq \min(1 - P_b, P_g) + 3|\Delta|. \]

The lower bound for \( P_\varepsilon \) can be obtained directly from decision theory [24].

Q.E.D.

REFERENCES


