Resilience of All-Optical Network Architectures under In-Band Crosstalk Attacks: A Graphical Model Approach[1]

Guanglei Liu and Chuanyi Ji

***Abstract*—An important question for secure all-optical networks (AONs) is how to incorporate security against attacks in the design and engineering of network architectures. In this work, we study the resilience of AON architectures under in-band crosstalk attacks. Crosstalk attack propagation depends on both optical devices at the physical layer and wavelength usage at the network layer. This motivates us to apply probabilistic graphical models to model attack propagation. At the physical layer, we use a directed probabilistic graph (Bayesian Belief Network) to model the attack propagation under static network traffic and a given source of attack. At the network layer, we use an undirected probabilistic graph to represent the probability distribution of active connections in the network. The cross-layer model is obtained by combining the physical- and the network-layer models into a factor graph representation. Graphical models provide an explicit representation of interactions between the physical- and the network layer. Furthermore, graphical models facilitate derivations of analytical results on resilience with respect to physical-layer vulnerability, physical topology, and network load. Specifically, we derive bounds on the network resilience for regular topologies. We show that for ring, star, and mesh-torus networks with link-shortest path routing and all-to-all traffic, the average network resilience loss grows linearly with respect to the network load when the network load is small, and polynomially with respect to the probability of attack propagation from node to node along the attacker's route. In addition, numerical results suggest that the sum-product algorithm based on the factor graph representation can be used for computationally efficient evaluation of network resilience for irregular/large topologies.**

Index Terms—Optical network architecture, security, network resilience, crosstalk attack, cross-layer, graphical models.

I. Introduction

All-optical network (AON) has been considered as a promising technology for next-generation optical networks. However, AONs are susceptible to malicious attacks as the signals remain in optical domain within the network and are difficult to be monitored closely [1]. Due to the high data rate supported by AONs, even attacks of a short duration can result in a large amount of data loss. Hence, security of AONs upon attacks has become an important issue, where an open question is how to incorporate security against attacks in the design and engineering of AON architectures. Investigations of this question are important as AONs are still at an early stage of implementation and ground-up developments of secure all-optical networks are possible [1]. The goal of this work is to study how network architecture impacts resilience in the context of in-band crosstalk attacks in AONs.

Crosstalk attacks were first studied in [1]. Crosstalk in AONs is caused by signal leakage among different inputs at non-ideal network devices, e.g. optical switches. The most detrimental type of crosstalk is in-band crosstalk, where the crosstalk element is within the same wavelength as the signal [3]. In-band crosstalk attacks can happen at fiber links or network nodes. In this work, we consider the case where an attacker gains legitimate access to a network node and inserts a flow with strong signal power into the network [2] [3]. Due to the crosstalk effects of wavelength switches, a small fraction of the signal from the attack channel may leak into other normal channels in the shared switching plane. The leakage superimposed onto normal channels may exceed a predetermined threshold for a quality of service requirement, such that those channels are considered to be affected by the attack at network nodes.

AONs are susceptible to crosstalk attacks. Major applications of AONs include metropolitan area networks (MANs) and wide area networks (WANs), but MANs and WANs are not 100% secure. As AON grows in span and functionality, it has the potential to provide services to a wider set of applications in the future, e.g. analog services, novel applications that require optical interfaces. Therefore, there is an increasing demand for access of the AON from outside parties, such as limited management access to the network from partners and customers of service providers, which results in an increasing threat to optical network security [4]. A wider set of users and an increasingly open platform of optical networks entail a higher risk of misuse of the network, which is evidenced by the security threats such as denial-of-service attack and worm attack in the current Internet [3]. It is expected that the risk of crosstalk attacks could be higher when the AON paradigm is fully implemented.

There have been several research activities aiming to mitigate the threats of crosstalk attacks in AONs. Attack detection based on node wrappers is studied in [3]. Necessary and sufficient conditions for crosstalk attack localization are investigated in [2]. General frameworks for managing faults and alarms in AON are discussed in [5-7]. All these approaches are reactive in nature. Furthermore, certain crosstalk attacks are difficult to detect [3]. For instance, sporadic crosstalk attacks may disrupt service but "disappear" before it can be detected. Thus, there remains a basic question: *How resilient is an AON upon crosstalk attacks before the attacks are detected and eliminated from the network?* This motivates our study of resilient AON architectures against crosstalk attacks. In a more general context, using cross-talk attacks as an example, we hope to provide an understanding

[1] Revised for *IEEE Journal on Selected Areas in Communications*, October 25, 2006. G. Liu and C. Ji are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250 USA (email: {guanglei, jic}@ece.gatech.edu). Research is supported in part by NSF grants ECS 0300605 and ECS 990857. An abridged version of this paper was presented at ISIT 2006, Seattle, USA, July 2006.

on how network architectures may affect network security in the presence of attackers.

We focus on three components of network architectures against crosstalk attacks: (a) physical layer optical devices, (b) physical topology, and (c) wavelength usage at the network layer, which is determined by network layer traffic. The goal is to quantify the effects of these factors against crosstalk attacks. One major challenge encountered in this study is to characterize the interactions of the three factors of network architecture during crosstalk attack propagation. For instance, attacks propagate to active wavelength channels of the same wavelength as the attacker's flow. Meanwhile, wavelength usage at the network layer is dependent because of the sharing of network links among different connections. Therefore, we need an approach that can provide an explicit representation of the cross-layer interactions.

We apply probabilistic graphical models to characterize cross-layer attack propagation [8]. Probabilistic graphical models include directed probabilistic graphs (Bayesian Belief Networks) and undirected probabilistic graphs (Markov Random Fields), and have been widely studied in machine learning and information theory [9] [10]. Yet, probabilistic graphical models have just begun to see applications in networking (see Section VIII for detailed discussions). In particular, at the physical layer, we develop a directed probabilistic graph to model attack propagation under static network traffic and a given source of attack with random attacking power. At the network layer, we apply an undirected probabilistic graph to represent the probability distribution of active connections. The physical- and the network-layer models together form a cross-layer model that has a factor graph representation [11].

The cross-layer model is developed using a bottom-up approach and provides an explicit representation of the complex dependencies between the physical- and the network-layer. Furthermore, the graphical models facilitate the analysis of multiple factors from network architecture on network resilience. For regular topologies, we derive bounds on the network resilience. For irregular/large topologies, the cross-layer model provides computationally efficient methods for studying the resilience where the analysis is not feasible.

The remainder of the paper is organized as follows. Section II describes the problem formulation. Section III presents the attack propagation model under static traffic and a given source of attack based on directed probabilistic graph. Section IV introduces the network-layer representation using undirected probabilistic graph. Section V discusses about the cross-layer model based on factor graph. Section VI investigates the impacts of physical layer on network resilience. Section VII studies the effects of the network layer on resilience. Section VIII briefly reviews graphical models in networking research. Section IX concludes the paper.

II. Problem Formulation

The topology of an AON is defined as an undirected graph $\mathrm{G}(\mathbf{V},\mathbf{E})$, with $\mathbf{V}$ being the set of nodes and $\mathbf{E}$ being the set of bi-directional links. Denote $V_i \sim V_j$ if there is one bi-directional link between V_i and V_j, $V_i, V_j \in \mathbf{V}$. Let $\mathbf{R}$ be a finite set of routes in the network. Assume that there are no wavelength converters in the AON. We define a connection on route r, $r \in \mathbf{R}$, as a bi-directional light-path on route r that consists of one unidirectional flow in each direction.[2] Assume that each wavelength can only be used by one active connection on the same network link.

This work considers single-source in-band crosstalk attacks. That is, crosstalk attack is started at the source node of a unidirectional flow on wavelength λ, and propagates to flows that use the same wavelength. As this work focuses on in-band crosstalk attacks, "flows", "connections'', and "channels" are used in the rest of the paper without referring to their associations with wavelength λ.

The problem we consider consists of two aspects: (a) developing the cross-layer model of attack propagation, and (b) using the model to quantify the network resilience upon crosstalk attacks.

Let S_i be a random variable that denotes the number of active channels affected by the in-band crosstalk attack at the switching plane of node V_i. Vector $\mathbf{S} = (S_i : V_i \in \mathbf{V})$ corresponds to the number of affected channels at each node in the network. Let N_{ij} denote the status of route r_{ij}, where $N_{ij} = 1$ if there is an active connection on route r_{ij} between node V_i and V_j, for $r_{ij} \in \mathbf{R}$; $N_{ij} = 0$, otherwise. Vector $\mathbf{N} = (N_{ij} : r_{ij} \in \mathbf{R})$ then represents the status of all network routes in $\mathbf{R}$. Denote f_{sd} as the flow starting from node s and terminating at node d. We need to obtain the following quantities to characterize attack propagation.

(a) $P(\mathbf{S} \mid \mathbf{N} = \mathbf{n},\ R_f = f_{sd})$: The probability of the number of channels affected at each network node given the status of network routes $\mathbf{n}$ and the source of attack R_f, where R_f denotes the unidirectional flow where the attack originates. This probability represents attack propagation under given $\mathbf{n}$ and f_{sd}, and is to be characterized through a directed probabilistic graph in Section III.

(b) $P(\mathbf{N} \mid R_f = f_{sd})$: The probability of the status of network routes given the source of attack, which is to be described using an undirected probabilistic graph in Section IV.

(c) $P(\mathbf{S} \mid R_f = f_{sd})$: The probability of the number of channels affected at each node given the source of attack, which models attack propagation under dynamic traffic. This probability combines the physical- and the network-layer models from (a) and (b), and shall be described with a factor graph representation in Section V.

[2] Each bi-directional link consists of two optical fibers, one for each direction. Throughout the paper, the term "connection" is used specifically for bi-directional traffic; the term "flow" is used to refer to uni-directional traffic.

The cross-layer model is then used to study network resilience based on the resilience loss for a given attack flow and the average resilience loss over all possible attack flows.

Definition 1: Given that there is a crosstalk attack started on flow f_{sd}, the network resilience loss is defined as

$$M_{f_{sd}} = \sum_{V_i \in \mathbf{V}} E_{f_{sd}}[S_i], \quad (1)$$

where $E_{f_{sd}}[S_i] = \sum_{S_i} s_i P(S_i = s_i \mid R_f = f_{sd})$ is the expected number of affected channel at node V_i given the source of the attack. $M_{f_{sd}}$ denotes the total number of active channels affected when the attack starts from a particular flow.

Definition 2: The average network resilience loss of the network is defined as $M = E_{R_f}[M_{f_{sd}}]$, where $E_{R_f}[\,]$ stands for the expectation over the source of the attack R_f, i.e.,

$$M = \sum_{f_{sd}} M_{f_{sd}} \; P(R_f = f_{sd}), \quad (2)$$

where

$$P(R_f = f_{sd}) = \frac{1}{2 \mid \mathbf{R} \mid} P(N_{sd} = 1), \quad (3)$$

with the assumption that each network route in **R** is equally likely to be an attacker's route, and the attack is started on one of the two unidirectional flows on the attacker's route with an equal probability.

III. Physical-Layer Attack Propagation Model: Directed Probabilistic Graph

We first model attack propagation under static network traffic and a given source of attack.

A. Background of In-Band Crosstalk Attack

We focus on in-band crosstalk attacks where an attacker gains legitimate access to the network and injects signals of high power into a flow. Due to imperfect switching arrays, the attacker's channel may affect other channels that share the switching plane, causing malfunctions at several locations in the network. Fig. 1 depicts an example of in-band crosstalk attack. At each network node, channels of the same wavelength from different input fibers share the same switching plane [12]. Suppose that the crosstalk is initiated on flow $C1$ using wavelength λ_1 from input fiber 1. All the wavelength channels that share a switching plane with $C1$, e.g. channel $C2$ from input 2, may be contaminated by $C1$'s power leakage.

In particular, we define a network node as being affected by the attack if the amount of in-band crosstalk incurred by normal[3] channels at the switching plane of that node exceeds a predetermined threshold. Clearly, each node along the attacker's route may be affected by the attack due to the high signal power of the attack flow, but the chance for nodes that are not on the attacker's route to be affected by the attack is negligible. That is, normal flows affected by the attack flow at one or more network nodes along the attacker's route do not have attacking capability, as its signal power is unlikely to be increased by more than half the normal channel power. For instance, consider the example in Fig. 1. Suppose, at one time instant, the attacker's jamming power is 20dB higher than the normal channel power and the optical switches have a crosstalk ratio of -35 dB. Then the power of flow C2 is increased by around -15dB of the normal channel power at node 1. The power of flow C2 at node 2 is in the same order as in node 1, whose crosstalk leakage to flow C3 is negligible given the crosstalk ratio of –35dB. Currently, optical switches with crosstalk ratios much less than -35dB are commercially available [13]. Thus, in this work, we ignore the in-band crosstalk caused by normal flows and assume that only nodes along the attacker's route may be affected by the attack. In addition, attack propagates to all the active channels that share the switching plane with the attacker's channel at each affected node. Based on these assumptions, we shall describe the probabilistic attack propagation model in the next subsection.

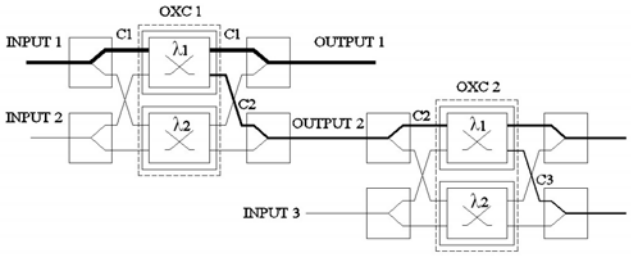

Fig. 1. Crosstalk attack propagation in AON.

B. Probabilistic Attack Propagation Model

Consider a crosstalk attack started at node s on flow f_{sd}. Let the set of nodes traversed by flow f_{sd} be $\mathbf{V}_{f_{sd}} = \{V_1, V_2, ..., V_k\}$, where V_1 and V_k are the source and the destination nodes respectively. The attack propagation is characterized by the status of each node in $\mathbf{V}_{f_{sd}}$ and the status of wavelength channels at the switches of those nodes.

We define the status of node V_i as a binary variable X_i. Specifically, let the signal power of a normal flow at the switching plane of each node be u_n when there is no attack in the network. Let the crosstalk ratio of the switches in the network be l_c and let c_{th} be a predetermined constant. Then $X_i = 1$ if the amount of in-band crosstalk incurred by a normal channel at the switching plane of node V_i exceeds $c_{th}u_n$; $X_i = 0$, otherwise. Furthermore, node V_i is affected by the attack if $X_i = 1$.

The use of binary nodal states facilitates the investigation of crosstalk attack propagation with random attacking power at the source. To be specific, the amount of in-band crosstalk at each node under attack may have a wide range of values. But the binary status of crosstalk levels at a node is simple and often observable. In fact, a frequent scenario for attack detection and monitoring is whether a predetermined threshold or service guarantee is violated at each node [5-7]. When the

[3] Normal channels refer to channels in the network that are not the attacker's channel.

amount of cross-talk is below the threshold, the node is "up", i.e., operational; otherwise, the node is "down", i.e., affected.

Hence we treat the attacker's jamming power as a random variable that obeys a certain probability distribution. The status of network nodes under crosstalk attacks then becomes binary random variables. The randomness lies in the fact that the cross-talk level is random due to the random jamming power of the attack. If the attacker's jamming power has a higher probability of being large; it is more likely for the attack to propagate farther away from the source node [7].

To determine the status of each node under attack, we now consider the attenuation of flow f_{sd} 's jamming power along its route. Denote the signal power of flow f_{sd} in the switching plane of node V_i as a random variable U_i, $i = 1, 2, ...k$. The attenuation of f_{sd} 's jamming power along its route can be captured using deterministic composite functions that depend on the characteristics of optical devices. For simplicity of illustration, we consider an example in Fig. 2.

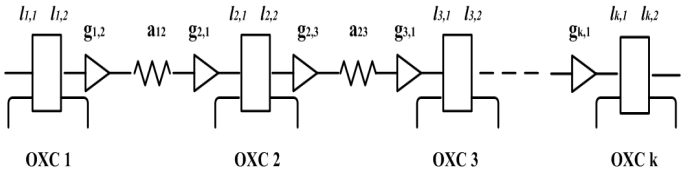

Fig. 2. Illustration of signal power attenuation: the attacker's flow.

We assume that there are input erbium-doped fiber amplifier (EDFA) and output EDFA at each side of a node respectively. Furthermore, we define the following parameters:

$l_{i,1}$: Signal loss ratio of node V_i before the flow enters the switching plane, which mainly includes signal loss at demultiplexer.

$l_{i,2}$: Signal loss ratio of node V_i after the flow enters the switching plane, which mainly includes loss at switching plane and multiplexer.

$a_{i,j}$: Signal loss ratio of the fiber span between node V_i and node V_j .

$g_{i,1}(\,)$: The gain of the EDFA at the input side of node V_i .

$g_{i,2}(\,)$: The gain of the EDFA at the output side of node V_i .

For a given network, $l_{i,1}, l_{i,2}, a_{i,j}$ are constants; $g_{i,1}(\,)$ and $g_{i,2}(\,)$ are deterministic non-linear functions of the input power to amplifiers. In this work, we adopt the following gain model for EDFAs [14]:

$$g_{ij}(P_{input}) = \begin{cases} d_{ij}, & if\ P_{input} \le p_{th}, \\ 1 + \dfrac{p_{sat}}{P_{input}} \lg \dfrac{g_0}{g_{ij}(P_{input})}, & otherwise, \end{cases} \quad (4)$$

where, P_{input} is the total input power; p_{sat} is the internal saturation power; g_0 is the small signal saturated gain; p_{th} is the input power threshold for successful gain clamping, and d_{ij} is the clamped gain value.

Assume that the attacker's flow (f_{sd}) does not share EDFAs with other flows. This corresponds to a conservative model of the jamming power attenuation and a worst-case scenario of in-band attack propagation, as all the photons of the EDFAs are used to amplify the attacker's signal. Then,

$$U_{i+1} = l_{i+1,1}\pi_{i+1,1}(a_{i,i+1}\pi_{i,2}(l_{i,2}U_i)), \quad (5)$$

where $\pi_{i,j} = P_{input} g_{i,j}(P_{input})$ is the output power of the EDFA with gain $g_{i,j}(P_{input})$ and input power P_{input} . Then composite function $\tau_{j-1,j}(\tau_{j-2,j-1}(...\tau_{i,i+1}(.))\,)$ capture the attenuation of the jamming power between node V_i and V_j .

Assume that, when there is no crosstalk attack in the network, amplifiers on each fiber operate in the gain clamped regions and make up the signal attenuation between two nodes. Furthermore, assume that the attacker's jamming power at the source node of the attack follows a cumulative distribution function $\eta(U)$ with minimum power $u_{\min}$, $u_{\min} \ge c_{th}u_n / l_c$, and maximum power $u_{\max}$. Then, it can be shown that the status of each node along the attacker's route, $X_i, i = 1, 2, ..., k,$ form a Markov Chain. Specifically,

$$P(X_1 = 1) = 1\,. \quad (6)$$

$$P(X_{i+1} \mid X_1, X_2, ..., X_i) = P(X_{i+1} \mid X_i), \quad i = 1, 2, ..., k-1. \quad (7)$$

$$P(X_{i+1} = 1 \mid X_i = 0) = 0. \quad (8)$$

$$P(X_{i+1} = 1 \mid X_i = 1) = \frac{P(U_{i+1} > c_{th}/l_c)}{P(U_i > c_{th}/l_c)} = \frac{1-\eta(\delta_{1,i+1})}{1-\eta(\delta_{1,i})}, \quad (9)$$

where $\delta_{1,i}, 1 \le i \le k-1,$ corresponds to the minimum value of jamming power at node V_1 such that attack can propagate to node V_i , and satisfies

$$\tau_{i-1,i}(\tau_{i-2,i-1}(...\tau_{1,2}(\delta_{1,i}))\,) = c_{th}u_n / l_c\,. \quad (10)$$

The derivation of (7) to (9) can be found in Appendix I.

The conditional probabilities in (9) can take different forms depending on $\eta(U)$. For simplicity, we can denote

$$P(X_{i+1} = 1 \mid X_i = 1) = \alpha_i\,, \quad (11)$$

where $\alpha_i = \dfrac{1-\eta(\delta_{1,i+1})}{1-\eta(\delta_{1,i})}$, where $\delta_{1,i}, 1 \le i < k$, as in (10).

If we further assume that the attacker's jamming power at the source node of the attack is uniformly distributed in $[u_{\min}, u_{\max}]$, (9) can be rewritten as

$$P(X_{i+1} = 1 \mid X_i = 1) = \frac{\max\{0, u_{\max} - \max\{u_{\min}, \delta_{1,i+1}\}\}}{u_{\max} - \max\{u_{\min}, \delta_{1,i}\}}. \quad (12)$$

In the rest of the paper, we assume that α_i 's are known.

We now consider the number of active channels affected by the attack at the switching plane of node V_i. Let $\mathbf{R}_{ij}$ be the set of network routes that use link ij. Under static traffic, $\sum_{r_{uv}\in\mathbf{R}_{ij}} n_{uv}$ corresponds to the number of flows that enter the switching plane of node V_i through link ij; $\sum_{r_{ih}\in\mathbf{R}_{ij}} n_{ih}$ corresponds to the number of flows that are locally originated from node V_i and enter the network through link ij. Hence, under static network traffic, the total number of affected channels at the switching plane of V_i, $V_i \in \mathbf{V}_{f_{sd}}$, is given by

$$P(S_i = s_i \mid X_i = x_i, \mathbf{N} = \mathbf{n}, R_f = f_{sd}) = \begin{cases} 1, & if\ s_i = \sum_{V_i \sim V_j} \{\sum_{r_{uv}\in\mathbf{R}_{ij}} n_{uv} + \sum_{r_{ih}\in\mathbf{R}_{ij}} n_{ih}\}\ \&\ x_i = 1, \\ & \quad or \\ & s_i = 1\ \&\ x_i = 0, \\ 0, & otherwise. \end{cases} \quad (13)$$

This means that, when node V_i is affected by the attack, all the active channels at the switching plane of V_i are affected by the attack; otherwise, if node V_i is not affected by the crosstalk attack, only one, i.e., only the channel used flow f_{sd} itself, is affected by the attack at the node.

Combining (11) and (13), we have the physical-layer attack propagation model,

$$P(\mathbf{S}_{f_{sd}} \mid \mathbf{N} = \mathbf{n},\ R_f = f_{sd}) = \sum_{\mathbf{X}_{f_{sd}}} \prod_{i=1}^{k-1} P(X_{i+1} \mid X_i, R_f = f_{sd}) \prod_{i=1}^{k} P(S_i \mid X_i, \mathbf{N} = \mathbf{n}, R_f = f_{sd}), \quad (14)$$

where $\mathbf{S}_{f_{sd}} = (S_i : V_i \in \mathbf{V}_{f_{sd}})$, and $\mathbf{X}_{f_{sd}} = (\mathbf{X}_i : V_i \in \mathbf{V}_{f_{sd}})$, which is the status of nodes in the attacker's route.

Therefore, under static network traffic (given $\mathbf{N} = \mathbf{n}$), $(X_i, S_i : V_i \in V_{f_{sd}})$ forms a directed probabilistic graph (Bayesian Belief Network). Each node in the probabilistic graph represents either X_i or S_i. There is one directed edge from X_i to X_{i+1} and one directed edge from X_i to S_i respectively. Note that, given $\mathbf{N} = \mathbf{n}$ and $X_i = x_i$, S_i is deterministic, but S_i is included for an explicit graphical representation of attack propagation.

Fig. 3 shows an example of a simple mesh network where all the routes in $\mathbf{R}$ are marked in dashed lines. Suppose that the crosstalk attack is started on flow *BD*. The directed probabilistic graph representation of attack propagation is shown in Fig. 4.

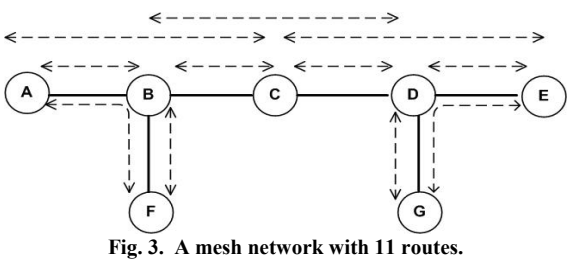

Fig. 3. A mesh network with 11 routes.

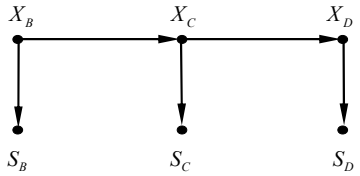

Fig. 4. Directed probabilistic graph representation of attack propagation: attack started on flow *BD*; mesh network in Fig. 3.

C. Discussions

Due to the complexity of modeling AON signal transmission, the physical-layer model in this section is developed with the following assumptions: (1) the in-band crosstalk due to channels with normal signal power and/or nonlinear effects is ignored; (2) under normal operations, the EDFAs work at gain-clamped region and make up for the signal losses between two network nodes; (3) the optical switches have the same crosstalk ratio and threshold of crosstalk leakage for the definition of node affection.

Assumption 1 is reasonable because of the low crosstalk ratio of current optical switches. If assumption 2 is relaxed so that the EDFAs work at gain-clamped region under normal operations, but may make up for more than the signal losses between two network nodes, then the status of nodes along the attacker's route may still form a Markov Chain. However, the order of nodes in the Markov Chain does not necessarily follow the sequence of $X_1, X_2, ..., X_k$. The same is true, if assumption 3 is relaxed so that optical switches in the network have different crosstalk ratios or the thresholds of crosstalk leakage for the definition of node affection are heterogeneous for different nodes.

Another approach to model the attack propagation is to define a random variable that corresponds to the position of the last affected node along the attacker's route. This model is equivalent to the Markov Chain model in (6) to (11), but does not include explicitly the status of each node. Thus such a model would not visually signify the actual attack propagation along the attacker's route.

The physical layer model characterizes attack propagation under static network traffic. Under dynamic traffic, however, the status of each network route N_{sd}, $r_{sd} \in \mathbf{R}$, is random and can be characterized using a network layer model.

IV. Network Layer Model: Undirected Probabilistic Graph

To obtain the network layer model, we need to obtain $P(\mathbf{N} \mid R_f = f_{sd})$, which is the probability distribution of route status given the source of the attack. From (3), we have

$$P(\mathbf{N} \mid R_f = f_{sd}) = P(\mathbf{N} \mid N_{sd} = 1). \quad (15)$$

Then it suffices to find $P(\mathbf{N})$, which can be characterized by undirected probabilistic graph.

A. Undirected Probabilistic Graph

An undirected probabilistic graph can be represented as **G = (V, E)** [8], where **V** represents the set of vertices, and

$\mathbf{E}$ represents the set of edges. Each node $V_i \in \mathbf{V}$ represents a random variable. A subset of nodes $\mathbf{V_C}$ is said to separate two other subsets of nodes $\mathbf{V_A}$ and $\mathbf{V_B}$ if every path joining every pair of nodes $V_i \in \mathbf{V_A}$ and $V_j \in \mathbf{V_B}$ has at least one node from $\mathbf{V_C}$ [8]. An undirected probabilistic graph implies a set of conditional independence relations. That is, for any disjoint subsets of nodes in the undirected graph, $\mathbf{V_A}$, $\mathbf{V_B}$, and $\mathbf{V_C}$, if $\mathbf{V_C}$ separates $\mathbf{V_A}$ and $\mathbf{V_B}$, then $\mathbf{V_A}$ and $\mathbf{V_B}$ are conditional independent given $\mathbf{V_C}$. For example, Fig. 5 shows an undirected probabilistic graph with 5 variables. As node V_2 and V_3 separate node V_1 from nodes in the rest of the network, $P(V_1 \mid V_2, V_3, V_4, V_5) = P(V_1 \mid V_2, V_3)$. Obviously, a node is separated from other nodes in the undirected graph by all its neighbors.

A clique denotes a subset of $\mathbf{V}$ that contains either a single node or several nodes which are all neighbors of one another. Then the joint probability distribution of $\mathbf{V}$ has a product form [8]:

$$P(\mathbf{V}) = Z^{-1} \prod_{q \in C} \psi_q(\{V_i : V_i \in \mathbf{V_q}\}), \quad (16)$$

where Z is the normalizing constant, $Z = \sum_{\mathbf{V}} \prod_{q \in C} \psi_q(\{V_i : V_i \in \mathbf{V_q}\})$; ψ_q is a non-negative function defined for clique $\mathbf{V_q} \in \mathbf{C}$, and $\mathbf{C}$ denotes the set of all the cliques in the graph $\mathbf{G}$.

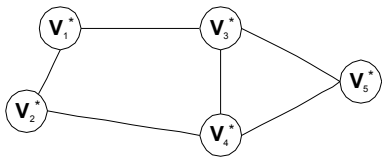

Fig. 5. An undirected probabilistic graph.

B. Network-Layer Model

The network-layer model is formed as follows. Each vertex in the undirected probabilistic graph represents the status of a route $N_{ij}, r_{ij} \in R$. Furthermore, the status of all network routes that share the same link forms a clique. In [15], it has been shown that the steady state distribution of the number of calls in loss networks without control form a Markov Random Field (MRF), which is one type of undirected probabilistic graph. Here we generalize the MRF representation in [15] to an undirected probabilistic graph representation that includes explicitly the dependence among different routes due to the capacity constraint and the network load.

We revisit the example of mesh network in Fig. 3, whose network-layer model is shown in Fig. 6. Consider route *AC*, which traverses two network links: *AB* and *BC*. Meanwhile, link *AB* is in route *AB* and route *AF*; link *BC* is in route *BC* and route *BD*. Since wavelength λ can only be used by one connection on each network link, route *AC* has a contention of wavelength usage with route *AB*, *AF*, *BC*, and *BD*. However, once the status of route *AB*, *AF*, *BC* and *BD* is known, the status of route *AC* can be determined without violating the capacity constraints. Hence, route *AB*, *AF*, *BC* and *BD* are neighbors of route *AC* and separate route *AC* from routes in the rest of the network, as shown in Fig. 6.

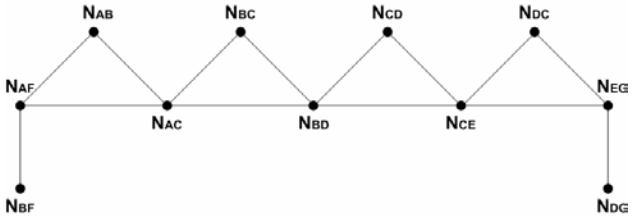

Fig. 6. Undirected probabilistic graph representation of network routes; mesh network in Fig. 3.

Therefore, by defining routes that share the same network link as neighbors, we capture the capacity constraint in the undirected probabilistic graph. The probability distribution of all network routes can be obtained by specifying proper clique potentials based on (16). The clique potentials in this work are selected to characterize both the dependencies among different network routes and the varying network load.

In Section III. B, we denote $\mathbf{R}_{ij}$ as a subset of routes in $\mathbf{R}$ that traverse link ij. A clique, denoted as C_{ij}, can then be formed with all the routes in $\mathbf{R}_{ij}$. Then the potential function of C_{ij}, denoted as ψ_{ij}, is obtained as follows: (1) $\psi_{ij} \neq 0$ if and only if the capacity constraint is satisfied, i.e., at most one route in $\mathbf{R}_{ij}$ is active; (2) if the wavelength is used on link ij, then $\psi_{ij} = \gamma_{ij}$; otherwise, $\psi_{ij} = 1 - \gamma_{ij}$, $0 < \gamma_{ij} < 1$. From (16), the joint probability of all routes satisfies

$$P(\mathbf{N}) = \frac{1}{Z_{\mathbf{N}}} \prod_{(V_i \sim V_j)} \gamma_{ij}^{\sum_{r_{uv} \in \mathbf{R}_{ij}} N_{uv}} (1-\gamma_{ij})^{(1-\sum_{r_{uv} \in \mathbf{R}_{ij}} N_{uv})} I_1(\sum_{r_{uv} \in \mathbf{R}_{ij}} N_{uv}), \quad (17)$$

where $I_1(\sum_{r_{uv} \in \mathbf{R}_{ij}} N_{uv}) = 1$ if $\sum_{r_{uv} \in \mathbf{R}_{ij}} N_{uv} = 0$ or 1; and $I_1(\sum_{r_{uv} \in \mathbf{R}_{ij}} N_{uv}) = 0$, otherwise. The clique functions are non-zero if and only if $I_1(\sum_{r_{uv} \in \mathbf{R}_{ij}} N_{uv}) = 1$. Thus (17) characterizes the dependencies of routes that result from the capacity constraints. Meanwhile, the network load, e.g. the probability that wavelength λ is used in the network, is characterized by parameters γ_{ij}'s. γ_{ij} can be considered as a "weight" for using a wavelength at link ij; $1-\gamma_{ij}$ can be considered as a "weight" for not using a wavelength at link ij. When $\gamma_{ij} \equiv \gamma, \forall V_i \sim V_j$, we can relate γ to the network load as follows:

Proposition 1: Let ρ denote the network load, $\rho = \boldsymbol{E}_{P(\mathbf{N})}[\frac{\sum_{V_i \sim V_j} \sum_{r_{uv} \in \mathbf{R}_{ij}} N_{uv}}{|\mathbf{E}|}]$. If in (17), $\gamma_{ij} \equiv \gamma, \forall V_i \sim V_j$, then ρ monotonically increases in γ.

Detailed proof of Proposition 1 can be found in Appendix II. Furthermore, (a) if $\gamma = 0.5$, the undirected probabilistic graph represents a uniform probability distribution on all

possible ways of using wavelength λ without violating the capacity constraint; (b) If $\gamma \to 1$, ρ increases toward the maximum value, which is determined by both the network topology and the route set $\mathbf{R}$; and (c) If $\gamma \to 0$, ρ approaches 0.

For simplicity of analysis, we assume that $\gamma_{ij} \equiv \gamma$, $\forall V_i \sim V_j$, in the rest of the paper. From (17), it follows that

$$P(\mathbf{N} \mid R_f = f_{sd}) \propto n_{sd} P(\mathbf{N} \setminus N_{sd}, N_{sd} = 1). \quad (18)$$

V. Cross-Layer Representation

The cross-layer model of attack propagation can be obtained by combining the physical- and the network-layer model using a factor graph [11], which corresponds to the following joint probability,

$$\begin{aligned} &P(\mathbf{X}_{f_{sd}}, \mathbf{S}_{f_{sd}}, \mathbf{N} \mid R_f = f_{sd}) \\ &= P(\mathbf{S}_{f_{sd}}, \mathbf{X}_{f_{sd}} \mid \mathbf{N}, R_f = f_{sd}) P(\mathbf{N} \mid R_f = f_{sd}), \end{aligned} \quad (19)$$

where $\mathbf{X}_{f_{sd}} = (X_i : V_i \in \mathbf{V}_{f_{sd}})$ and $\mathbf{S}_{f_{sd}} = (S_i : V_i \in \mathbf{V}_{f_{sd}})$.

Factor graph is a bipartite graph showing how a global function can be factorized into a product of local functions. Each local function depends on a subset of the variables [11]. There are two types of nodes in a factor graph: a variable node for each variable, and a factor node for each local function. There is an edge connecting a variable node to a factor node if and only if the variable is an argument of the local function.

Fig. 7 shows the factor graph representation for the mesh network in Fig. 3 when the attack is started from flow *BD*. The lower portion of the factor graph represents attack propagation at the physical layer. As the attack may propagate from node V_i to V_{i+1}, $V_i, V_{i+1} \in \mathbf{V}_{f_{sd}}$, X_i and X_{i+1} are connected to the same factor node $P(X_{i+1} \mid X_i, R_f = f_{BD})$. Furthermore, the number of affected channels at node V_i is determined by X_i and routes that traverses node V_i. Therefore, S_i, X_i, and those routes passing through node V_i are connected to the factor node that describes the conditional probability in (13).

The upper portion of the factor graph characterizes the dependence at the network layer. All the network routes that share a common network link ij are connected to the clique function ψ_{ij} in (17). Here, the factor graph provides an explicit representation of the dependencies among different network components during attack propagation.

Factor graphs subsume directed and undirected probabilistic graphical models, and provide explicit representations of the factorization of probability distributions [11]. The application of factor graph provides two advantages: (1) It shows the intricate dependencies among different network components during a crosstalk attack; (2) it provides computationally efficient algorithms to evaluate the network resilience loss, which shall be discussed in Section VII.

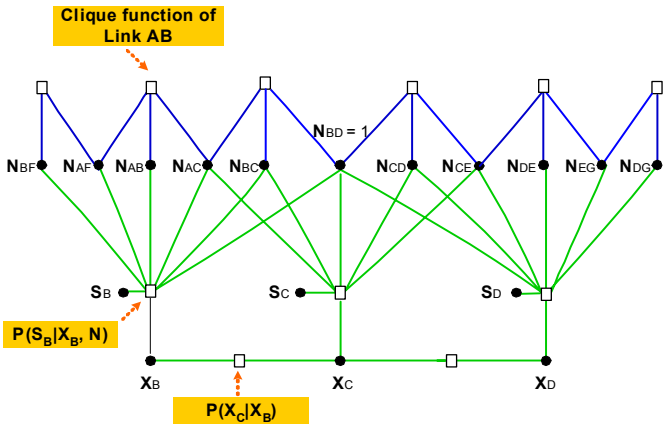

Fig. 7. Factor graph representation of mesh network in Fig. 3; attack is started on flow *BD*.

VI. Network Resilience: Impact of Physical Layer

We now use the cross-layer model to study the network resilience. We begin with the physical layer and quantify how the resilience varies with physical topology as well as the physical layer vulnerabilities, characterized by α_i in (11).

A. Impacts of Physical-Layer Vulnerabilities on Network Resilience Loss

We consider the impact of physical-layer vulnerabilities by considering the lower and upper bounds of network resilience loss $M_{f_{sd}}$. The lower bound of $M_{f_{sd}}$ results from the best-case scenario of resilience upon attack: there is no active connections on wavelength λ that traverses link ij, $\forall V_i \in \mathbf{V}_{f_{sd}}$, $V_j \notin \mathbf{V}_{f_{sd}}, V_i \sim V_j$. In this case, at the switching plane of each node along the attacker's route, only two channels are active that correspond to the connection on the attacker's route. The upper bound result from the worst-case scenario of network resilience upon attack: there always exists an active connection inserted into the network at node V_i and traverses link ij, $\forall V_i \in \mathbf{V}_{f_{sd}}$, $V_j \notin \mathbf{V}_{f_{sd}}, V_i \sim V_j$. In this case, the number of active channels in the switching plane of node V_i is $2(d_i - 1)$ or $2d_i$, $\forall V_i \in \mathbf{V}_{f_{sd}}$, where d_i is the nodal degree of V_i.

For simplicity, we assume $\alpha_i \equiv \alpha, \forall V_i \in \mathbf{V}_{f_{sd}}$. Then, the network resilience loss can be bounded as in the following proposition.

Proposition 2: The network resilience loss for a given source of attack f_{sd} can be bounded as

$$k + \sum_{i=1}^{k} \alpha^{i-1} \le M_{f_{sd}} \le k + 2(1 + \alpha^{k-1}) + \sum_{i=1}^{k} (2d_i - 3)\alpha^{i-1}, \quad (20)$$

where k is the total number of nodes in $\mathbf{V}_{f_{sd}}$, $k > 1$.

The lower bound in (20) characterizes the effect of route-length and α on attack propagation, which increases polynomially with respect to α. Furthermore,

$$k+\sum_{i=1}^{k}\alpha^{i-1} = \begin{cases} k+1+\alpha+o(\alpha), & as\ \alpha\to 0, \\ 2k-0.5k(k-1)(1-\alpha)+o(1-\alpha), & as\ \alpha\to 1, \end{cases} \quad (21)$$

which shows that $M_{f_{sd}}$ is determined by the length of route r_{sd}. The upper bound in (20) increases approximately linearly with α, where

$$k+2(1+\alpha^{k-1})+\sum_{i=1}^{k}(2d_i-3)\alpha^{i-1} = \begin{cases} k+2d_1-1+(2d_2-3)\alpha+o(\alpha), & as\ \alpha\to 0\ \ and\ \ k>2, \\ k+2d_1-1+(2d_2-1)\alpha+o(\alpha), & as\ \ \alpha\to 0\ and\ \ k=2, \\ (4-2k)+2\sum_{i=1}^{k}d_i \\ \quad +\{\frac{3}{2}k^2-\frac{7}{2}k+2-\sum_{i=1}^{k}2d_i(i-1)\}(1-\alpha)+o(1-\alpha), & as\ \alpha\to 1, \end{cases} \quad (22)$$

which shows that, when there is always an active connection inserted into the network at node V_i using link ij, $\forall V_i \in \mathbf{V}_{f_{sd}}, V_j \notin \mathbf{V}_{f_{sd}}, V_i \sim V_j$, if the network vulnerability is low, $M_{f_{sd}}$ is determined by the route length and the nodal degree of the source node of the attack. If the network vulnerability is high, $M_{f_{sd}}$ is determined by the total number of network links incidental on nodes along the attacker's flow, i.e., the number of links in set $\mathbf{E}_{f_{sd}} = \{e_{ij} : V_i \in \mathbf{V}_{f_{sd}}\}$. In addition, $|\mathbf{E}_{f_{sd}}| = \sum_{i=1}^{k} d_i + (1-k)$.

B. Impact of Physical Topology on Network Resilience Loss

We use the lower- and upper-bound in (20) to study the impact of physical topology on $M_{f_{sd}}$. For clarity, we summarize the asymptotic results on $M_{f_{sd}}$ in (21) and (22) for network resilience under various topologies in Table I. Assume that there is one link-shortest route between each pair of nodes in the network. The asymptotic properties of these topologies are summarized in Table II ([16][17]). Combining the impacts of physical-layer vulnerability and physical topology, we find that,

(1) If the physical-layer vulnerability is high ($\alpha \to 1$),

(i) The upper bound of $M_{f_{sd}}$ shows that fully-connected mesh network and star network are the least resilient due to the large size of the set $\mathbf{E}_{f_{sd}}$.

(ii) The lower bound of $M_{f_{sd}}$ shows that network with a ring topology is generally the least resilient because of the large route length in a ring network.

(2) If the physical-layer vulnerability is low ($\alpha \to 0$),

(i) The upper bound of $M_{f_{sd}}$ shows that the fully-connected mesh topology is the least resilient since each node in the network has nodal degree $m-1$.

(ii) The lower bound of $M_{f_{sd}}$ shows that the ring network is generally the least resilient due to the large route length.

(3) Chord networks exhibit good resilience whose resilience loss $M_{f_{sd}}$ increases logarithmically with respect to the number of nodes in the network in the worst case.

Note that in addition to the resilience measure considered in this work, there exists other performance metrics for network resilience, e.g. two-terminal connectivity [18], and flexibility in route selection [16]. Therefore, different performance metrics of network resilience need to be considered simultaneously when choosing a resilient network design. Overall, a chord network offers excellent resilience upon crosstalk attacks and good route selection flexibility.

Table I
Bounds of Network Resilience Loss $M_{f_{sd}}$

α bounds	Upper bound of $M_{f_{sd}}$	Lower bound of $M_{f_{sd}}$
$\alpha\to 1$	$2\sum_{i=1}^{k}d_i+(4-2k)-O(1-\alpha)$	$2k-O(1-\alpha)$
$\alpha\to 0$	$k+2d_1-1+O(\alpha)$	$k+1+O(\alpha)$

Table II
Asymptotic Properties of Different Network Topologies with *m* Nodes

Topology	Ave. nodal degree	Ave. route length	Ave. size of $\mathbf{E}_{f_{sd}}$
Star	1	2	m
Ring	2	$m/4$	$m/4$
n-ary Tree	m+1	$O(\log_n m)$	$O(\log_n m)$
Mesh-Torus	4	$O(\sqrt{m})$	$O(\sqrt{m})$
Fully-Connected Mesh	m	1	m
Chord [23]	$\log_2 m$	$O(\log_2 m)$	$O(\log_2 m)$

VII. NETWORK RESILIENCE: IMPACT OF NETWORK LAYER

We now study the impact of network layer on the resilience in terms of network load. In particular, we are interested in quantifying how the network resilience varies jointly with the load, and the physical-layer vulnerability α.

A. Impact of Network Load on Network Resilience

We first consider the impact of network load ρ on $M_{f_{sd}}$. From (1),

$$M_{f_{sd}} = \sum_{V_i \in \mathbf{V}_{f_{sd}}} \boldsymbol{E}_{f_{sd}}[S_i], \quad (23)$$

where $\boldsymbol{E}_{f_{sd}}[S_i]$ is the mean number of channels affected by the attack at the switching plane of node $V_i, V_i \in \mathbf{V}_{f_{sd}}$.

Furthermore,

$$E_{f_{sd}}[S_i] = 1+\alpha^{i-1}\{1+\sum_{V_i\sim V_j, V_j\notin \mathbf{V}_{f_{sd}}}\{E_{f_{sd}}[\sum_{r_{uv}\in \mathbf{R}_{ij}} N_{uv}]+E_{f_{sd}}[\sum_{r_{ih}\in \mathbf{R}_{ij}} N_{ih}]\}\}, \quad \forall V_i \in \mathbf{V}_{f_{sd}}, \quad (24)$$

where $E_{f_{sd}}[\sum_{r_{ih}\in \mathbf{R}_{ij}} N_{ih}]$ is the mean number of active channels that are locally inserted into the network at node V_i and leave node V_i through link ij, and $E_{f_{sd}}[\sum_{r_{uv}\in \mathbf{R}_{ij}} N_{uv}]$ is the mean

number of active flows that enter node V_i through link ji, given that the attack starts from flow f_{sd}.

Does the network resilience loss $M_{f_{sd}}$ always increase with ρ for arbitrary network with arbitrary route? The answer is no. For a counterexample, please refer to [19]. However, when practical route sets are considered, $M_{f_{sd}}$ increases with ρ for several typical network topologies.

Theorem 1: For a *ring* network, assume the route set $\mathbf{R}$ consists of the two-link disjoint routes between each pair of nodes in the network. Let m be the number of nodes traversed by the attacker's flow f_{sd}. Then, $M_{f_{sd}}$ monotonically increases in ρ. In particular, $M_{f_{sd}}$ satisfies

$$v_1 + 2(1+\alpha^{m-1})\gamma \le M_{f_{sd}} \le v_1 + 2(1+\alpha^{m-1})\rho, \quad (25)$$

where $v_1 = m + 2\sum_{i=1}^{m}\alpha^{i-1}$. Furthermore, for $0 < \rho << 1$, $\rho = \gamma + o(\rho)$, and the upper and the lower bounds meet

$$M_{f_{sd}} = v_1 + 2(1+\alpha^{m-1})\rho + o(\rho). \quad (26)$$

Detailed proof of Theorem 1 can be found in Appendix III.

Theorem 2: For a *star* network, assume that the route set $\mathbf{R}$ consists of the routes between each pair of nodes in the network. Let m, $m>1$, be the number of nodes in the network. Let the hub node be denoted as V_m. Then, $M_{f_{sd}}$ monotonically increases in ρ. In particular, $M_{f_{sd}}$ satisfies

$$M_{f_{sd}} \ge \begin{cases} 3+\alpha+(m-2)\alpha\gamma, & if\ f_{sd} = f_{A_iA_m}, i=1,...,m-1, \\ 3+\alpha+(m-2)\gamma, & if\ f_{sd} = f_{A_mA_i}, i=1,...,m-1, \\ 4+\alpha+\alpha^2+(m-3)\alpha\gamma, & otherwise, \end{cases} \quad (27)$$

$$M_{f_{sd}} \le \begin{cases} 3+\alpha+2(m-2)\alpha\rho, & if\ f_{sd} = f_{A_iA_m}, i=1,...,m-1, \\ 3+\alpha+2(m-2)\rho, & if\ f_{sd} = f_{A_mA_i}, i=1,...,m-1, \\ 4+\alpha+\alpha^2+2(m-3)\alpha\rho, & otherwise. \end{cases} \quad (28)$$

Furthermore, for $0 < \rho << 1$, the bounds are tight, and

$$M_{f_{sd}} = \begin{cases} 3+\alpha+2(m-2)\alpha\rho+o(\rho), & if\ f_{sd} = f_{A_iA_m}, i=1,...,m-1, \\ 3+\alpha+2(m-2)\rho+o(\rho), & if\ f_{sd} = f_{A_mA_i}, i=1,...,m-1, \\ 4+\alpha+\alpha^2+2(m-3)\alpha\rho+o(\rho), & otherwise, \end{cases} \quad (29)$$

Proofs of Theorem 2 can be found in Appendix IV. $M_{f_{sd}}$ generally is the sum of two terms: e.g. in (30), $(3+\alpha)$ that corresponds to the number of affected channels used by flows on the attacker's route; and $(2(m-2)\alpha\rho+o(\rho))$ that corresponds to the number of affected channels used by flows not on the attacker's route.

We also compare $M_{f_{sd}}$ for ring and star networks. In both cases, $M_{f_{sd}}$ is linearly increasing in ρ for $0 < \rho << 1$. However, for ring network, $M_{f_{sd}}$ is polynomially increasing in α; whereas for star network, $M_{f_{sd}}$ is linearly increasing in α. For ring networks, $M_{f_{sd}}$ is linearly increasing in k (the number of nodes in $\mathbf{V}_{f_{sd}}$). For star networks, $M_{f_{sd}}$ is linearly increasing in m (the number of nodes in the network).

B. Impact of Network Load on Average Resilience Loss

We now focus on the impact of network load ρ on the average network resilience loss (M), which is the mean value of network resilience loss over all possible source of attacks.

Consider a ring network with m, $m>1$, nodes, $V_1, V_2, ..., V_m$, and a route set $\mathbf{R}$ that includes all the two link-disjoint paths between each pair of nodes in the network. Then, we have

Theorem 3: $$M_{ring,m} = \frac{1}{m-1}\sum_{i=1}^{m-1} a_i M_{f_{1\,i+1}}, \quad (30)$$

where $a_i = P(N_{1\,i+1} = 1)$ is the probability that a connection with i links between two terminal nodes, V_1 and V_{i+1}, is active, and $M_{f_{1\,i+1}}$ is the network resilience loss when the attack is started from flow $f_{1\,i+1}$. Furthermore,

$$a_i = \theta^i f_{m-i+1}/g_m, \quad (31)$$

$$\theta = \gamma/(1-\gamma), \quad (32)$$

$$f_m = \frac{\sqrt{1+4\theta^2}+1}{2\sqrt{1+4\theta^2}}\left(\frac{1+2\theta+\sqrt{1+4\theta^2}}{2}\right)^{m-1} + \frac{\sqrt{1+4\theta^2}-1}{2\sqrt{1+4\theta^2}}\left(\frac{1+2\lambda-\sqrt{1+4\theta^2}}{2}\right)^{m-1}, \quad (33)$$

$$g_m = f_m + \sum_{j=1}^{m-1} j\theta^j f_{m+1-j}, \quad m>1. \quad (35)$$

Detailed proof of Theorem 3 can be found in Appendix V. Using Theorem 1, we have the following bounds

$$M_{ring,m} \ge \frac{1}{(m-1)}\sum_{i=1}^{m-1}\{a_i(i+1+\sum_{j=0}^{i}\alpha^j + 2(1+\alpha^i)\gamma)\}, \quad (34)$$

$$M_{ring,m} \le \frac{1}{(m-1)}\sum_{i=1}^{m-1}\{a_i(i+1+\sum_{j=0}^{i}\alpha^j + 2(1+\alpha^i)\rho)\}, \quad (35)$$

The difference between the upper and the lower bound of $M_{ring,m}$ is $O((\rho-\gamma)/m)$. Furthermore, (30) can be simplified as

$$M_{ring,m} = \rho M_{f_{A_1A_2}}/(m-1) + o(\rho), \text{ as } \rho \to 0. \quad (36)$$

$$M_{ring,m} = \sum_{i=1}^{m-1}\frac{1}{2^{m+1}} M_{f_{A_1A_{i+1}}}/(m-1), \text{ as } \rho \to 1, m \to \infty. \quad (37)$$

Then, we have

$$M_{ring,m} = \rho(3+\alpha)/(m-1) + o(\rho), \text{ as } \rho \to 0, \quad (38)$$

which shows that, when the network load is low, $M_{ring,m}$ increases almost linearly with ρ and α; and is in the order of $O(\rho/m)$. When the network load is high,

$$M_{ring,m} = \frac{1}{m-1}\{\sum_{i=1}^{m-1}\frac{1}{2^i}(i+1+\sum_{j=0}^{i}\alpha^j + 2(1+\alpha^i)\}, \quad (39)$$

as $\rho \to 1, m \to \infty$. Furthermore, if $\alpha = 1$, (39) can be simplified as

$$M_{ring,m} = \frac{1}{m-1}(6 - \frac{m+5}{2^{m-1}}), \quad (40)$$

which shows that $M_{ring,m}$ is in the order of $O(1/m)$.

Next consider a star network with m nodes, $V_1, V_2, ..., V_m$, where V_m is the hub node of the star network; and a route set $\mathbf{R}$ that includes all the link-disjoint paths between each pair of nodes in the network. It follows that,

Theorem 4:

$$M_{star,m} = \frac{b_1(M_{f_{A_1A_k}} + M_{f_{A_kA_1}}) + 2b_2(m-2)M_{f_{A_1A_2}}}{2(m-1)}, m > 3, \quad (41)$$

where $b_1 = P(N_{A_1A_k} = 1)$, $b_2 = P(N_{A_1A_2} = 1)$ with $b_1 = \theta t_{m-1}/t_m$; $b_2 = \theta^2 t_{m-2}/t_m$; $t_1 = 1$; $t_2 = 1+\theta$; $t_i = (1+\theta)t_{i-1} + (i-2)\theta^2 t_{i-2}$, $\forall\, i > 2$.

Detailed proof of Theorem 4 can be found in Appendix VI. Furthermore, using Theorem 2, we have the following bounds for $M_{star,m}$,

$$\begin{aligned} M_{star,m} \geq \frac{1}{2(m-1)}\{&b_1(4 + (1+\alpha)(2 + (m-2)\gamma)) \\ &+ 2b_2(m-2)(4 + \alpha + \alpha^2 + (m-3)\alpha\gamma)\}, \end{aligned} \quad (42)$$

$$\begin{aligned} M_{star,m} \leq \frac{1}{2(m-1)}\{&b_1(4 + 2(1+\alpha)(1 + (m-2)\rho)) \\ &+ 2b_2(m-2)(4 + \alpha + \alpha^2 + 2(m-3)\alpha\rho)\}. \end{aligned} \quad (43)$$

The difference between the upper and the lower bound of $M_{star,m}$ is $O(\alpha(2\rho - \gamma))$ when m is large, since b_2 is $O(1/m)$. In addition, when the network load is low,

$$M_{star,m} = \rho(3+\alpha)/m + o(\rho), \text{ as } \rho \to 0, \quad (44)$$

which shows that $M_{star,m}$ is $O(\rho/m)$; and increases linearly with α. When the network load is high ($\rho \to 1$) , we have

$$M_{star,m} = O(\alpha\rho), \text{ as } \rho \to 1, \quad (45)$$

which shows that, when the star network is under high load, $M_{star,m}$ increases linearly with α.

For a general network $\boldsymbol{G}(\mathbf{V},\mathbf{E})$ with a fixed set of route $\mathbf{R}$, we have the following upper bound for M:

Theorem 5: $$M \leq \frac{1}{|\mathbf{R}|}\max_{f_{sd}}\{\mathbf{M}_{f_{sd}}\}\rho|\mathbf{E}|, \quad (46)$$

where $|\mathbf{E}|$ is the cardinality of the set of edges in the network; $|\mathbf{R}|$ is the cardinality of the route set $\mathbf{R}$.

The proof of Theorem 5 can be found in Appendix VII. In (46), $\rho|\mathbf{E}|/|\mathbf{R}|$ corresponds to the upper bound of the probability that the crosstalk attack occurs in the network, and is accurate when the network route set $\mathbf{R}$ only consists of routes with link-length 1. The bound in (46) provides a worst case estimation of M. Furthermore, suppose that all the routes in the set $\mathbf{R}$ are of the same link length l. Then the probability that an crosstalk attack happens in the network is $(\rho|\mathbf{E}|)/(l|\mathbf{R}|)$, and (46) can be refined as

$$M \leq \frac{1}{|\mathbf{R}|\,l}\max_{f_{sd}}\{M_{f_{sd}}\}\rho|\mathbf{E}|. \quad (47)$$

When the length of each network route and the network resilience loss $M_{f_{sd}}$ are the same for each possible source of attack, the equality in (47) holds. Theorem 5 suggests the upper bound of average network resilience loss is affected by the following factors:

(1) The network load in the network. The upper bound in (47) increases at least linearly with ρ.

(2) The number of links in the network. The larger the number of links in the network, the less resilient the network. The upper bound in (47) increases linearly with the number of links in the network.

(3) The number of routes in the network. The larger the number of routes in the network, the more resilient the network. This is because that the probability for a route to be chosen as the attacker's route is smaller.

Next we use (46) to study a mesh-torus network with m nodes and a route set $\mathbf{R}$, which includes: (1) the unique link-shortest route between each pair of nodes if applicable; and (2) one shortest route between each pair of nodes, which forms the border of the sub-grid with the two nodes at the diagonally opposite corners.

Theorem 6:

$$\max_{f_{sd}}\{M_{f_{sd}}\} \leq \begin{cases} \frac{6(1-\alpha^{\sqrt{m}+1})}{(1-\alpha)} + 4, & \text{if } \alpha \neq 1, \\ 6\sqrt{m} + 4, & \text{otherwise.} \end{cases} \quad (48)$$

Then, from (46), we have $M_{torus,m} \leq \frac{2m\rho}{m(m-1)}\max_{f_{sd}}\{M_{f_{sd}}\}$,

$$M_{torus,m} \leq \begin{cases} \frac{2\rho}{(m-1)}(\frac{6(1-\alpha^{\sqrt{m}+1})}{(1-\alpha)} + 4), & \text{if } \alpha \neq 1, \\ \frac{2\rho}{(m-1)}(6\sqrt{m} + 4), & \text{otherwise.} \end{cases} \quad (49)$$

Furthermore, when $\rho \to 0$, it can be found that

$$M_{torus,m} = \rho(3+\alpha)/(m-1) + o(\rho), \quad \text{as } \rho \to 0. \quad (50)$$

Detailed proof of Theorem 6 is omitted here.

We compare the average network resilience loss for ring, star and mesh networks in Table III. It can be observed that:

(1) When the network load is low $(0 < \rho << 1)$, M is $O(\rho/m)$. This is because when the load is close to 0, the network is most likely in either of two states: (a) there is no active connection in the network; or (b) there is an active connection of link length 1. Specifically, with probability $O(\rho)$, the attack is started on a route of link length 1; with

probability $o(\rho)$, the attack is started on a route of longer lengths. For instance, as each route is the attacker's route with equal probability, the attack starts on routes of link length 1 in the mesh-torus network with probability $2\rho/(m-1)$ and $M_{f_{sd}} = (3+\alpha)+o(\rho)$ if $\rho << 1$.

(2) When the network load is high ($\rho \to 1$), the star network is the least resilient, with M being $O(\alpha)$. This is because, for the star network, nodes in the set $\mathbf{V}_{f_{sd}}, \forall r_{sd} \in \mathbf{R}$, has the most number of neighboring links. Ring and mesh-torus networks show good network resilience in $O(\rho/m)$.

Table III
Average Network Resilience Loss (*M*)

M	$\rho \to 0$	$\rho \to 1$
Ring network	$\rho(3+\alpha)/(m-1)$	$O(1/m)$
Star network	$\rho(3+\alpha)/m$	$O(\alpha)$
Mesh-torus	$2\rho(3+\alpha)/(m-1)$	$\begin{cases} O(1/(1-\alpha)m), & \text{if } \alpha \neq 1, \\ O(1/\sqrt{m}), & \text{otherwise.} \end{cases}$

C. Irregular Topologies

For networks with irregular topologies, we resort to the sum-product algorithm on the factor graph. The sum-product algorithm is then compared with the exact resilience calculation through enumerations of all network traffic patterns. Enumeration has the computational complexity exponential in the number of routes, and is thus not applicable to networks with even a medium number of routes. The computational complexity of the sum-product algorithm is exponential in the maximum nodal degree of the factor graph for the worst case [11], and is thus much more efficient than enumeration. The sum-product algorithm provides exact results when the factor graph has no loops, and provides approximate results otherwise [11].

When there are a large number routes in the set **R**, to further reduce the computational complexity of the sum-product algorithm, the following intermediate variables can be introduced: (1) $W_{ij} = \sum_{r_{uv} \in \mathbf{R}_{ij}} N_{uv}, W_{ij} \in \{0,1\}$, which is the number of flows that enter the switching plane of node V_i through link ij; (2) $H_{ij} = \sum_{r_{ih} \in \mathbf{R}_{ij}} N_{ih}, H_{ij} \in \{0,1\}$, which is the number of flows locally originated at node i and leaves node V_i through link ij. Then the factor graph representation can be transformed accordingly. On the other hand, it is possible to transform factor graphs with loops into loop-free factor graphs, so that exact results can be obtained using the sum-product algorithm at the cost of computational complexity [11].

We first consider three networks shown in Fig. 8. In each network, the route set has 21 routes, which corresponds to one link-shortest route between each pair of nodes. Using the sum-product algorithm, we first compute the network resilience loss given the source of attack $M_{f_{sd}}$ for each f_{sd}. We then use the sum-product algorithm to find the probability of $P(N_{sd}=1)$. Finally, (2) is used to compute the average network resilience loss.

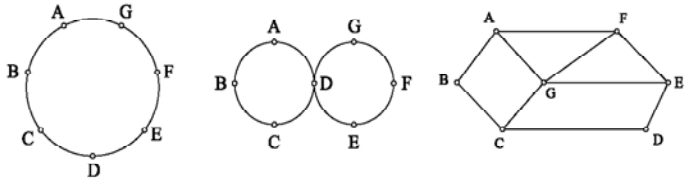

Fig. 8. Ring, double-ring, and mesh networks.

Fig. 9 depicts the relationship between ρ and average network resilience loss M for the networks in Fig. 8. with $\alpha = 0.6$. It can be observed that:

(1) M monotonically increases with ρ, in networks with all-to-all traffic and link-shortest path routing. Moreover, for low load, M increases linearly with ρ.

(2) The sum-product algorithm results in an almost exact M for the mesh and ring networks, even though the factor graph representations contain loops. The performance of sum-product algorithm is not as accurate yet acceptable for the double-ring network. This suggests that the sum-product algorithm can be used for large networks where exact calculation of resilience is infeasible.

We use the sum-product algorithm to study the network resilience for the NSF network topology [20] with 14 nodes and 21 bi-directional links. Assume that there is one link-shortest route between each pair of nodes in the networks. Then, there are 91 routes in **R**. The corresponding factor graph representation contains loops, and thus sum-product algorithm provides an approximation for M.

Fig. 10 shows the relationship between ρ and M for the NSF network topology with $\alpha = 0.3, 0.6, 0.9$. It suggests that, if the set of network routes consists of one link-shortest route between each pair of nodes in the network, M generally increases with the network load. Furthermore, when the network load is low, M increases linearly with ρ.

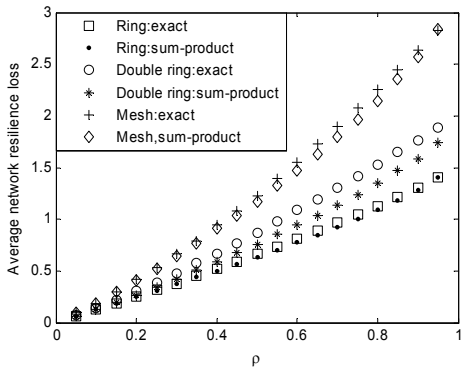

Fig. 9. Average network resilience loss vs. network load; $\alpha = 0.6$, three networks in Fig. 8.

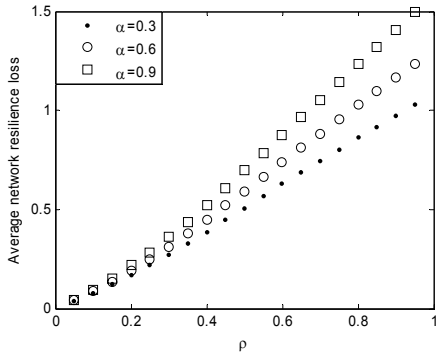

Fig. 10. Average network resilience loss vs. network load; $\alpha = 0.3, 0.6, 0.9$; NSF network topology.

VIII. Related Work in Probabilistic Graphical Models

Bayesian Belief Network has been used in fault localization and detection (see [21] and references therein) in complex communication systems. There the construction of the Bayesian Belief Network representation is usually assumed based on the causal relationships and the main difficulty then lies in estimation of model parameters, i.e., conditional probabilities. In [22], dependency graph is used for IP fault localization, where the problem of shared-risk-link-group is formulated using a bipartite graph.

Markov Random Field (MRF) has been used to study the blocking probability in a loss network [15] as discussed in Section IV. This work generalizes the work in [15] by using undirected probabilistic graphs to capture the basic dependency among different routes due to the capacity constraint, and the network load. In [23], self-localization in sensor networks is formulated using MRF with single-node and pair-wise clique potentials, and Belief Propagation algorithm is used for self-calibration.

Factor graphs have been applied to the problem of multicast link loss inference in sensor networks [24]. The factor graph is constructed based on the concept of link and route costs with two fundamental assumptions: link costs are assumed to mutually independent and path flows are assumed to be mutually independent. In this work, we maintain the dependence among different flows, and obtain the factor graph through decomposition of the joint probability distribution of network nodes, links and connections. In [25], factor graph has been used for scalable source/channel decoding in large-scale sensor networks. In this case, factor graph provides a simplified model of the correlation among sensor data and enables scalable iterative decoding.

In [26], a similar bottom-up approach is used to derive a cross-layer model for self-configuration of ad hoc wireless networks based on probabilistic graphical models. The resulting model is shown to be a Markov Random Field for the physical- and the link layer [26].

IX. Conclusions

In this work, we have studied resilience of all-optical networks (AONs) under in-band crosstalk attacks. Our goal is to develop a cross-layer model that characterizes attack propagation in the network, and to study the resilience of AON architectures from both the physical- and the network layer. We have found that probabilistic graphical models can serve this goal. In particular, at the physical layer, a directed probabilistic graph can model attack propagation when network under static traffic. At the network layer, an undirected probabilistic graph can represent the probability distribution of active connections. A cross-layer model is then obtained by combining the physical- and network-layer models into a factor graph representation.

There are several benefits resulting from the cross-layer model based on graphical models. The model provides an explicit representation of the dependencies and interactions between the physical- and the network layer. In addition, it facilitates the analytical investigation of network resilience for ring, star, and special cases of mesh topologies. Finally, the cross-layer model facilitates the implementation of computationally efficient approaches, e.g. the sum-product algorithm, for evaluating network resilience.

Through both analysis and numerical study, we have explored several factors from both the physical- and the network layer that affect the resilience. Factors from the physical-layer include: (1) the physical-layer vulnerability, parameters in Bayesian Belief Network that characterize how likely the attack propagates, and (2) the physical topology. Factors from the network layer include active network connections that are characterized using network load, i.e., the probability that the wavelength, on which the attack is initiated, is used in the network. We show that for all the topologies studied in this paper, the average network resilience loss increases linearly with respect to the physical-layer vulnerability and light network load under link-shortest routing, and all-to-all traffic. In addition, ring and mesh-torus network show good resilience, which are inversely proportional to the number of the nodes in the network. Numerical results also suggest that for networks with link-shortest routing and all-to-all traffic, the network resilience loss increases at least linearly with respect to the network load.

There are several open issues for future research. One is how to develop physical-layer models of attack propagation under less stringent conditions. Another is to consider the effect of dynamic network resource allocation algorithms upon attack propagation. Finally, the technique of probabilistic graphical models to other failure/attack problems in all-optical networks.

Acknowledgment

The authors would like to acknowledge valuable comments from anonymous reviewers, helpful discussions with Ali Adibi on physical layer models of attack propagation, and helpful suggestions from Sung-Eok Jeon, Zesheng Chen, Rajesh Narasimha, and Supaporn Erjongmanee.

References

[1] M. Medard, D. Marquis, R. A. Barry, and S. G. Finn, "Security issues in all-optical networks," *IEEE Network*, Vol. 11, No. 3, pp. 42-48, May/June 1997.

[2] T. Wu and A. K. Somani, "Cross-talk attack monitoring and localization in all-optical network," *IEEE/ACM Transactions on Networking,* Vol. 13, No. 6, pp. 1390-1401, December 2005.

[3] M. Medard, S. R. Chinn, and P. Saengudomlert, "Node wrappers for QoS monitoring in transparent optical nodes," *Journal of High Speed Networks*, Vol. 10, pp. 247-268, 2001.

[4] "Secure Optical Networks," Cisco Systems, *White Paper*, Copyright 1999-2005.

[5] J. K. Patel, S. U. Kim, D. H. Su, S. Subramaniam, and H-A Choi, "A framework for managing faults and attacks in WDM optical networks," Proc. of the *DAPRA* Information Survivability Conference and Exposition, 2001.

[6] S. Stanic, S. Subramaniam, H. Choi, G. Sahin, and H-A Choi, "Efficient alarm management in optical network," Proc. of the *DAPRA* Information Survivability Conference and Exposition, 2003.

[7] C. M. Machuca, I. Tomkos, and O. K. Tonguz, "Failure location algorithm for transparent optical networks," *IEEE Journal on Selected Areas in Communications*, Vol. 23, pp. 1508-1519, Aug. 2005.

[8] P. Smyth, D. Heckerman, and M. I. Jordan, "Probabilistic independence networks for hidden markov probability models," *Neural Computation*, 9, pp. 227-270, 1997.

[9] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter, "Probabilistic networks and expert systems," Springer-Verlag, New York, 1999.

[10] S. Geman and K. Kochanek, "Dynamic programming and the graphical representation of error-correcting codes," *IEEE Transactions on Information Theory*, 47, pp. 549-568, 2001.

[11] F. R. Kschischang, B. J. Frey, and H-A Loeliger, "Factor graph and the sum-product algorithm," *IEEE Transactions on Information Theory*, Vol. 47, No. 2, pp. 498-519, 2001.

[12] J. Zhou, R. Cadeddu, E. Casaccia, C. Cavazzoni, and M. J. O'Mahony, "Crosstalk in multiwavelength optical cross-connect networks," *IEEE Journal of Lightwave Technology*, Vol. 12, pp. 1423-1435, June 1996

[13] A Tzanakaki, I. Zacharopoulos, and I. Tomkos, "Broadband building blocks," *IEEE Circuits and Devices Magazine*, pp. 32-37, Mar./Apr. 2004.

[14] T. Deng and S. Subramaniam, "An analysis of optical amplifier gain competition attack in a point-to-point WDM link," *Proc. of Opticomm*, pp. 249-261, July 2002.

[15] S. Zachary and Ilze Ziedins, "Loss networks and Markov Random Field," *Journal of Applied Probability*, No. 2, pp. 403-414, 1999.

[16] K. Gummadi, R. Gummadi, S. Gribble, S. Ratnasamy, S. Shenker, and I. Stoica, "The impact of DHT routing geometry on resilience and proximity," Proc. of *SIGCOMM 2003*, Karlsruhe, Germany.

[17] Behrooz Parhami, "Introduction to parallel processing: algorithms and architectures," Plenum Press, 1999.

[18] G. Weichenberg, V. Chan, and M. Médard, "High-reliability architectures for networks under stress," *IEEE Journal on Selected Areas in Communications*, Vol. 22, pp. 1830-1845, 2005.

[19] G. Liu and C. Ji, "Graphical Models for Resilience of All-Optical Networks under In-Band Crosstalk Attacks," *Technical Report*, http://users.ece.gatech.edu/guanglei/jsac.html, School of ECE, Georgia Insitute of Technology, 2006.

[20] H. Zhu, H. Zang, K. Zhu, and B. Mukherjee, "A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks," *IEEE/ACM Transactions on Networking,* Vol. 11, No. 2, pp. 285-299, Apr. 2003.

[21] M. Steinder and A. S. Sethi, "Probabilistic fault localization in communication systems using belief networks," *IEEE Transactions on Networking*, No. 5, pp. 809-822, Oct. 2004.

[22] R. R. Kompella, J. Yates, A. Greenberg, and A. C. Snoeren, "IP fault localization via risk modeling," Proceedings of the 2nd ACM/USENIX Symposium on Networked Systems Design and Implementation (NSDI) Boston, MA, May 2005.

[23] A. T. Ihler, J. W. Fisher III, R. L. Moses, and A. S. Willsky, "Nonparametric belief propagation for self-localization of sensor networks," *IEEE Journal on Selected Areas in Communications*, Vol. 23, pp. 809-819, April 2005.

[24] Y. Mao, F. R. Kschischang, B. Li, and S. Pasupathy, "A factor graph approach to link loss monitoring in wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, Vol. 23, pp. 820-829, April 2005.

[25] J. Barros, M. Tuchler, and S. P. Lee, "Scalable source/channel decoding for large-scale sensor networks," Proc. of *IEEE* International Conference on Communications, 2004.

[26] Sung-eok Jeon and Chuanyi Ji, "Nearly optimal distributed configuration management using probabilistic graphical model," Proc. of Workshop on Resource Provisioning and Management in Sensor Networks, pp. 219-226, MASS 2005.

Appendix I: Derivation of (7) to (9)

Let U_i denote the jamming power of the attack flow at node $V_i, \forall V_i \in \mathbf{V}_{f_{sd}}$. We first show that if $U_i \geq U_{i+1}$ for $1 \leq i < k$, then $P(X_{i+1} \mid X_1, X_2 ..., X_i) = P(X_{i+1} \mid X_i)$.

Suppose $U_i \geq U_{i+1}$. Since $X_i = 1$, if $U_i > c_{th}/(l_c u_n)$; and $X_i = 0$, otherwise; it follows that $P(X_1 = x_1, X_2 = x_2, ..., X_i = x_i) \neq 0$, only if $x_1 \geq x_2 \geq ... \geq x_i$. Let $k_1 = \max\{j : x_j = 1 \,\&\, 1 \leq j \leq i\}$, which is the largest index of nodes affected by the attack among $V_1, V_2, ..., V_i$. Then,

(a) If $1 \leq k_1 < i$, $X_1 = 1, ... X_{k_1} = 1, X_{k_1+1} = 0, ..., X_i = 0$.

$$\begin{aligned} &P(X_{i+1} = 0 \mid X_1 = 1, ... X_{k_1} = 1, X_{k_1+1} = 0, ..., X_i = 0) \\ &= \frac{P(X_1 = 1, ..., X_{k_1} = 1, X_{k_1+1} = 0, ..., X_i = 0, X_{i+1} = 0)}{P(X_1 = 1, ..., X_{k_1} = 1, X_{k_1+1} = 0, ..., X_i = 0)} \\ &= \frac{P(X_{k_1} = 1, X_{k_1+1} = 0)}{P(X_{k_1} = 1, X_{k_1+1} = 0)} = 1. \end{aligned} \tag{51}$$

Obviously, $P(X_{i+1} = 0 \mid X_i = 0) = 1$. Therefore, $P(X_{i+1} \mid X_1, X_2 ..., X_i) = P(X_{i+1} \mid X_i)$.

(b) If $1 \leq k_1 = i$, $X_1 = 1, ..., X_i = 1$.

$$\begin{aligned} &P(X_{i+1} = 0 \mid X_1 = 1, ..., X_i = 1) \\ &= \frac{P(X_1 = 1, ..., X_i = 1, X_{i+1} = 0)}{P(X_1 = 1, ..., X_i = 1)} \\ &= \frac{P(U_1 > c_{th}/(l_c u_n), ..., U_i > c_{th}/(l_c u_n), U_{i+1} < c_{th}/(l_c u_n))}{P(U_1 > c_{th}/(l_c u_n), ..., U_i > c_{th}/(l_c u_n))} \\ &= \frac{P(U_i > c_{th}/(l_c u_n), U_{i+1} < c_{th}/(l_c u_n))}{P(U_i > c_{th}/(l_c u_n))} \\ &= \frac{P(X_i = 1, X_{i+1} = 0)}{P(X_i = 1)} \\ &= P(X_{i+1} = 0 \mid X_i = 1). \end{aligned} \tag{52}$$

Next we show $U_i \geq U_{i+1}$, assuming that, when there is no crosstalk attack in the network, amplifiers on each fiber operate in the gain clamped regions and make up the signal attenuation between the two nodes. From (5), we have

$$U_{i+1} = \tau_{i,i+1}(U_i) = l_{i+1,1}\pi_{i+1,1}(a_{i,i+1}\pi_{i,2}(l_{i,2}U_i)).$$

and,

$$l_{i+1,1}a_{i,i+1}l_{i,2}d_{i,2}d_{i+1,1} = 1, \tag{53}$$

where $d_{i,2}$ denotes the clamped gain of EDFA at the output side of node V_i; $d_{i+1,1}$ denotes the clamped gain of EDFA at the input side of node V_{i+1}. Then,

$$\text{If } l_{i,2}U_i \le p_{th,(i,2)},\ U_{i+1} = U_i; \tag{54}$$

$$\text{If } l_{i,2}U_i > p_{th,(i,2)},\ \pi_{i,2}(l_{i,2}U_i) < d_{i,2}U_i, \tag{55}$$

which corresponds to the case where the EDFA with subscript $(i,2)$ works at the saturation region. Therefore,

$$U_{i+1} < l_{i+1,1}d_{i+1,1}a_{i,i+1}d_{i,2}l_{i,2}U_i.$$

It follows that $U_{i+1} < U_i$. To prove (9), it suffices to show that $\tau_{i,\,i+1}(U_i)$ monotonically increases in U_i, where

$$U_{i+1} = \tau_{i,i+1}(U_i) = l_{i+1,1}\pi_{i+1,1}(a_{i,i+1}\pi_{i,2}(l_{i,2}U_i)).$$

Since $l_{i+1,1}, a_{i,i+1},$ and $l_{i,2}$ are constants, to show that $\tau_{i,\,i+1}(U_i)$ monotonically increases in U_i, it suffices to show that $\pi_{ij}(P_{input})$ monotonically increase in P_{input}. This can be obtained by showing $\dfrac{\partial(P_{input}g(P_{input}))}{\partial P_{input}} > 0$ for (4). This means that the higher the input power at EDFA, the higher the output power, when the EDFA worked at either the saturated or the non-saturated region. Detailed calculation of the derivative is omitted.

Appendix II: Proof of Proposition 1

To prove Proposition 1, it suffices to show that

$$\frac{\partial\rho}{\partial\gamma},\ \forall\ 0<\gamma<1.$$

From (17), assume $\gamma_{ij} \equiv \gamma$, then,

$$P(\mathbf{N}) = \frac{1}{Z_{\mathbf{N}}}\prod_{(V_i\sim V_j)}\gamma^{\sum_{r_{sd}\in\mathbf{R}_{ij}}N_{sd}}(1-\gamma)^{(1-\sum_{r_{sd}\in\mathbf{R}_{ij}}N_{sd})}I_1(\sum_{r_{sd}\in\mathbf{R}_{ij}}N_{sd}). \tag{56}$$

Let, $W_{ij} = \sum_{r_{uv}\in\mathbf{R}_{ij}} N_{uv}$ and $\mathbf{W} = (W_{ij} : V_i \sim V_j)$. $\mathbf{W}$ is a vector that represent the wavelength usage at each link in the network. We denote a configuration of $(\mathbf{N},\mathbf{W})$ with non-zero probability as a traffic pattern, i.e., a traffic pattern $(\mathbf{N},\mathbf{W})$ satisfies the capacity constraints and $P(\mathbf{N}=\mathbf{n},\mathbf{W}=\mathbf{w})>0$. Let $\mathbf{T}_k, k=0,1,...,|\mathbf{E}|$, be the set of traffic patterns that k links in the network are used by active connections, with $|\mathbf{E}|$ being the number of links in $\mathbf{E}$. Let $|\mathbf{T}_k|$ denote the cardinality of $\mathbf{T}_k$, then,

$$\begin{aligned}\rho &= E_{P(\mathbf{N})}[\textstyle\sum_{V_i\sim V_j}\sum_{r_{sd}\in\mathbf{R}_{ij}}N_{sd}\Big/|\mathbf{E}|]\\ &= \frac{\sum_{k=0}^{|\mathbf{E}|}k\gamma^k(1-\gamma)^{|\mathbf{E}|-k}\,|\mathbf{T}_k|}{|\mathbf{E}|\sum_{k=0}^{|\mathbf{E}|}\gamma^k(1-\gamma)^{|\mathbf{E}|-k}\,|\mathbf{T}_k|}\\ &= \frac{\sum_{k=1}^{|\mathbf{E}|}k\theta^k\,|\mathbf{T}_k|}{|\mathbf{E}|\sum_{k=0}^{|\mathbf{E}|}\theta^k\,|\mathbf{T}_k|} \overset{definition}{=} \xi(\theta),\end{aligned} \tag{57}$$

where $\theta = \gamma/(1-\gamma), \theta>0$.

$$\begin{aligned}&\frac{\partial(\xi(\theta))}{\partial\theta}\\ &= \frac{1}{|\mathbf{E}|(\sum_{k=0}^{|\mathbf{E}|}\theta^k\,|\mathbf{T}_k|)^2}\cdot\{(\textstyle\sum_{k=1}^{|\mathbf{E}|}k^2\theta^{k-1}\,|\mathbf{T}_k|)(\sum_{k=0}^{|\mathbf{E}|}\theta^k\,|\mathbf{T}_k|)\\ &\qquad -(\textstyle\sum_{k=1}^{|\mathbf{E}|}k\theta^k\,|\mathbf{T}_k|)(\sum_{k=1}^{|\mathbf{E}|}k\theta^{k-1}\,|\mathbf{T}_k|)\}.\end{aligned} \tag{58}$$

Using Cauchy-Schwartz Inequality, it can be shown that

$$\partial(\xi(\theta))/\partial\theta > 0, \forall\ \theta > 0. \tag{59}$$

Since $\partial\theta/\partial\gamma > 0 \ \forall\ 0<\gamma<1$, we have

$$\partial\rho/\partial\gamma > 0, \forall\ 0<\gamma<1. \tag{60}$$

Therefore, ρ increases monotonically in γ.

Appendix III: Proof of Theorem 1

Consider a ring network $G(\mathbf{V},\mathbf{E})$ with m nodes $(m>1)$. The route set $\mathbf{R}$ consists of the two link-disjoint routes between each pair of nodes in the network. Suppose the crosstalk attack is started on flow f_{ij} between node V_i and V_j, $i,j=1,2,...,m,\ i<j$. The set of nodes traversed by flow f_{ij} is $\mathbf{V}_{f_{ij}} = \{V_i, V_{i+1},...,V_j\}$. Then at most two nodes, node V_{i-1} and node V_{j+1}, are neighbors of nodes in $\mathbf{V}_{f_{ij}}$, but are not in $\mathbf{V}_{f_{ij}}$ themselves. Without loss of generality, we focus on the conditional wavelength usage at link between V_j and V_{j+1}.

To show that $M_{f_{ij}}$ monotonically increases in ρ for the ring network, from (24), it suffices to show that $E_{f_{ij}}[\sum_{r_{uv}\in\mathbf{R}_{j\,j+1}}N_{uv}]$ and $E_{f_{ij}}[\sum_{r_{jh}\in\mathbf{R}_{j\,j+1}}N_{jh}]$ monotonically increase with parameter γ in (56) for the ring network. Let $W_{ij} = \sum_{r_{uv}\in\mathbf{R}_{ij}}N_{uv}$ and $H_{ij} = \sum_{r_{ih}\in\mathbf{R}_{ij}}N_{ih}$. Then for the ring network, denote $E_{f_{ij}}[\sum_{r_{uv}\in\mathbf{R}_{j\,j+1}}N_{uv}]$ as $w_{j\,j+1}(m,f_{ij},ring)$; denote $E_{f_{ij}}[\sum_{r_{jh}\in\mathbf{R}_{j\,j+1}}N_{jh}]$ as $\varpi_{j\,j+1}(m,f_{ij},ring)$, where m is the number of nodes in the ring network.

Let $w_{12}(l,bus)$ denote the mean value of W_{12} in an l-node network of bus topology with a route set that includes the route between each pair of nodes, where subscript l denotes the number of nodes in the bus network. Since,

$$w_{j\,j+1}(m,f_{ij},ring) = \varpi_{j\,j+1}(m,f_{ij},ring) = w_{12}(m-j+i,bus), \tag{61}$$

it is sufficient to show that $w_{12}(l,bus), \forall\ l>1,$ increases monotonically with γ.

As in Appendix II, let $\theta = \gamma/(1-\gamma)$ and $\mathbf{W} = (W_{ij} : V_i \sim V_j)$. In addition, a configuration of $(\mathbf{N},\mathbf{W})$ with non-zero probability is denoted as a traffic pattern. Let $sum(\mathbf{W})$ denote the summation of all the components in $\mathbf{W}$, then from (56),

$$P(\mathbf{N},\mathbf{W}) \propto \theta^{sum(\mathbf{W})}\prod_{V_i\sim V_j}I_2(W_{ij} = \sum_{r_{uv}\in\mathbf{R}_{ij}}N_{uv}), \tag{62}$$

where $I_2(A)=1$ if A is true; and $I_2(A)=0$, otherwise. If $(\mathbf{N},\mathbf{W})$ is a traffic pattern, (62) can be simplified as

$$P(\mathbf{N},\mathbf{W}) \propto \theta^{sum(\mathbf{W})} . \quad (63)$$

Let $\mathbf{T}_{(l),bus}$ denote the set of all traffic patterns on the bus network with l nodes ($l>1$). By counting all possible ways of using link A_1A_2, we have for the l-node bus network,

$$P(W_{12}=1) \propto \theta \sum_{\mathbf{T}_{(l-1),bus}} \theta^{sum(\mathbf{W}_{(l-1),bus})} + \theta^2 \sum_{\mathbf{T}_{(l-2),bus}} \theta^{sum(\mathbf{W}_{(l-2),bus})} + ... + \theta^l;$$

$$P(W_{12}=0) \propto \sum_{\mathbf{T}_{(l-1),bus}} \theta^{sum(\mathbf{W}_{(k-1),bus})} . \quad (64)$$

Let $f_l(\theta) = \sum_{\mathbf{T}_{(l),bus}} \theta^{sum(\mathbf{W}_{(l),bus})}$, $\forall l>1$, and $f_1(\theta)=1$. We have,

$$P(W_{12}=1) \propto \theta f_{l-1}(\theta) + \theta^2 f_{l-2}(\theta) + ... + \theta^l ,$$

$$P(W_{12}=0) \propto f_{l-1}(\theta) .$$

Furthermore, we have the following recursive equations,

$$f_1(\theta)=1;\ f_2(\theta)=1+\theta;$$

$$f_i(\theta) = (1+2\theta) f_{i-1}(\theta) - \theta f_{i-2}(\theta),\ \ i=3,4,...,l. \quad (65)$$

Then,

$$w_{12}(l,bus) = \begin{cases} \dfrac{\theta}{1+\theta}, & if\ l=2, \\ 1 - \dfrac{f_{l-1}(\theta)}{(1+2\theta) f_{l-1}(\theta) - \theta f_{l-2}(\theta)}, & if\ l>2. \end{cases} \quad (66)$$

Obviously, $\dfrac{\partial w_{12}(l,bus)}{\partial \theta} > 0$, for $l=2$. If $l>2$,

$$\frac{\partial w_{12}(l,bus)}{\partial \theta} = \frac{2f_{l-1}^{'}(\theta) + \theta(f_{l-1}^{'} f_{l-2} - f_{l-1} f_{l-2}^{'})}{((1+2\theta) f_{l-1}(\theta) - \theta f_{l-2}(\theta))^2} . \quad (67)$$

$$f_l^{'} f_{l-1} - f_l f_{l-1}^{'} = 2f_{l-1}^2 - f_{l-1} f_{l-2} + \theta(f_{l-1}^{'} f_{l-2} - f_{l-1} f_{l-2}^{'}) . \quad (68)$$

Since $f_2^{'} f_1 - f_2 f_1^{'} = 2\theta^2 + 4\theta + 1 > 0$, through Mathematical induction, from (68), we have $f_l^{'} f_{l-1} - f_l f_{l-1}^{'} > 0, \forall l>1$, and

$$\frac{\partial w_{12}(l,bus)}{\partial \theta} > 0,\ \forall\ l>1 . \quad (69)$$

Since $\theta = \gamma/(1-\gamma)$, and $0<\theta<1$, it follows that $\dfrac{\partial w_{12}(l,bus)}{\partial \gamma} > 0, \forall l>1$. From Proposition 1, $M_{f_{ij}}$ monotonically increases in ρ for the ring network.

The upper and lower bound of $M_{f_{sd}}$ in (25) is obtained by showing that

$$\gamma \le w_{j\,j+1}(m, f_{ij}, ring) \le \rho . \quad (70)$$

Here we first show that for arbitrary network topology $\mathbf{G}(\mathbf{V},\mathbf{E})$, if $r_{ij} \in \mathbf{R}$, i.e., there is one route from node V_i to V_j in $\mathbf{R}$, where $V_i \in \mathbf{V}_{f_{sd}}$, $V_j \notin \mathbf{V}_{f_{sd}}$, and $V_i \sim V_j$, then,

$$E_{f_{sd}}[W_{ij}] \ge \gamma , \quad (71)$$

Obviously, the ring network considered here satisfies the condition in (71).

Let $\mathbf{E}_1$ be the set of links that are not traversed by flow f_{sd}: $\mathbf{E}_1 = \{e_{ij} : V_i \sim V_j, r_{sd} \notin \mathbf{R}_{ij}\}$. Let $\mathbf{R}_1$ be the set of routes in $\mathbf{R}$ that only traverse links in $\mathbf{E}_1$. Let $\mathbf{E}_2 = \mathbf{E}_1 \setminus \{e_{ij}\}$, and $\mathbf{R}_2$ be the set of routes in $\mathbf{R}$ that only traverse links in $\mathbf{E}_2$. Clearly, $\mathbf{R}_2 \subset \mathbf{R}_1 \subset \mathbf{R}$, if $r_{ij} \in \mathbf{R}$.

Let $\mathbf{T}_{\mathbf{E}_1} = \{(\mathbf{N}_{\mathbf{E}_1}, \mathbf{W}_{\mathbf{E}_1})\}$ be the set of traffic patterns restricted to a network formed by link set $\mathbf{E}_1$ with route set $\mathbf{R}_1$. Let $\mathbf{T}_{\mathbf{E}_2} = \{(\mathbf{N}_{\mathbf{E}_2}, \mathbf{W}_{\mathbf{E}_2})\}$ be the set of traffic patterns restricted to a network formed by link set $\mathbf{E}_2$ with route set $\mathbf{R}_2$. Then,

$$E_{f_{sd}}[W_{ij}] = \frac{\theta Z_1(\theta) + Z_2(\theta)}{(1+\theta) Z_1(\theta) + Z_2(\theta)}, \quad (72)$$

where $Z_1(\theta) = \sum_{\mathbf{T}_{\mathbf{E}_2}} \theta^{sum(\mathbf{W}_{\mathbf{E}_2})}$,

$Z_2(\theta) = \sum_{\mathbf{T}_{\mathbf{E}_1}} \theta^{sum(\mathbf{W}_{\mathbf{E}_1})} - (1+\theta) Z_1(\theta)$. In addition, $Z_2(\theta) > 0$ if there is a route that traverses link ij and one or more links in set $\mathbf{E}_2$; $Z_2(\theta) = 0$, otherwise. Since $Z_1(\theta) > 0$, we have

$$E_{f_{sd}}[W_{ij}] \ge \frac{\theta}{1+\theta} = \gamma . \quad (73)$$

To show that $w_{j\,j+1}(m, f_{ij}, ring) \le \rho$, from (61), it suffices to show that

$$w_{12}(m-j+1, bus) \le \rho , \quad (74)$$

which can be proved through the following two lemmas.

Lemma 1: $w_{12}(l,bus) \le w_{12}(l+1,bus),\ \forall\ l>1.$

Lemma 2: $w_{12}(m,bus) \le \rho, \forall m>1.$

Lemma 1 and 2 are proved using induction similarly as in the proof of (69). Detailed proof is omitted here. Using (70), we can obtain (25) from (24).

Appendix IV: Proof of Theorem 2

For a network of star topology with m nodes, $m>2$, and a route set $\mathbf{R}$ that consists of the routes between each pair of nodes. Let node V_m be the hub node of the star network. Let $W_{ij} = \sum_{r_{iv} \in \mathbf{R}_{ij}} N_{uv}$ and $H_{ij} = \sum_{r_{ih} \in \mathbf{R}_{ij}} N_{ih}$. We first show that, when the attack is started on flow f_{1m}, $w_{mi}(m, f_{1m}, star) + \varpi_{mi}(m, f_{1m}, star)$, $i=2,...,m-1$, increases monotonically with γ, where

$$w_{mi}(m, f_{1m}, star) = E_{f_{1m}}[\sum_{r_{uv} \in \mathbf{R}_{mi}} N_{uv}], \text{ and}$$

$$\varpi_{mi}(m, f_{1m}, star) = E_{f_{1m}}[\sum_{r_{mh} \in \mathbf{R}_{mi}} N_{mh}].$$

Let $\mathbf{T}_{(l),star}$ denote the set of all traffic patterns on the star network with l nodes. By counting all possible ways of using link mi, it can be found that, for the l-node star network,

$$P(W_{mi}=1 \mid R_f = f_{1m}) \propto \theta \sum_{\mathbf{T}_{(m-1),star}} \theta^{sum(\mathbf{W}_{(m-1),star})} + (m-2)\theta^2 \sum_{\mathbf{T}_{(m-2),star}} \theta^{sum(\mathbf{W}_{(m-2),star})}; \quad (75)$$

$$P(W_{mi}=0 \mid R_f = f_{1m}) \propto \sum_{\mathbf{T}_{(m-1),star}} \theta^{sum(\mathbf{W}_{(m-1),star})} ; \quad (76)$$

$$P(H_{mi}=1 \mid R_f = f_{1m}) \propto \theta \sum_{\mathbf{T}_{(m-1),star}} \theta^{sum(\mathbf{W}_{(m-1),star})} ; \quad (77)$$

$$P(H_{mi}=0 \mid R_f=f_{1m})$$
$$\propto \sum_{\mathbf{T}_{(m-1),star}} \theta^{sum(\mathbf{W}_{(m-1),star})} + (m-2)\theta^2 \sum_{\mathbf{T}_{(m-2),star}} \theta^{sum(\mathbf{W}_{(m-2),star})}. \quad (78)$$

Let $t_1=1$ and $t_l(\theta)=\sum_{\mathbf{T}_{(l),star}} \theta^{sum(\mathbf{W}_{(l),star})}$, $l>1$. Then, we have the following recursive equations,

$t_2=1+\theta;\ t_l=(1+\theta)t_{l-1}+(l-2)\theta^2 t_{k-2},\ \forall\, l>2$.

Therefore, from (75)- (78),

$$w_{mi}(m,f_{1m},star)+\varpi_{mi}(m,f_{1m},star)$$
$$=1+\frac{(\theta-1)t_{m-1}}{(1+\theta)t_{m-1}+(m-2)\theta^2 t_{m-2}}. \quad (79)$$

Through induction similarly as in the proof of (69), we have $\frac{\partial\{\, w_{mi}(m,f_{1m},star)+\varpi_{mi}(m,f_{1m},star)\}}{\partial\gamma}>0$. Similarly, when the attack is started from flow $f_{A_1A_2}$, it can be shown that $\frac{\partial\{w_{mi}(m,f_{12},star)+\varpi_{mi}(m,f_{12},star)\}}{\partial\gamma}>0$. Thus, from (24), it follows that $M_{f_{ij}}$ monotonically increases in ρ for the star network. The upper and lower bound of $M_{f_{sd}}$ in (27) and (28) is obtained by showing that

$$\gamma < w_{mi}(m,f_{1m},star)+\varpi_{mi}(m,f_{1m},star)\le 2\rho,$$
$$\gamma < w_{mi}(m,f_{12},star)+\varpi_{mi}(m,f_{12},star)\le 2\rho. \quad (80)$$

Since

$$\varpi_{mi}(m,f_{1m},star)\le w_{mi}(m,f_{1m},star), \text{ and}$$
$$\varpi_{mi}(m,f_{12},star)\le w_{mi}(m,f_{12},star), \quad (81)$$

Equation (80) can be obtained by showing

$$\gamma \le w_{mi}(m,f_{1m},star)\le \rho,$$
$$\gamma \le w_{mi}(m,f_{12},star)\le \rho. \quad (82)$$

The proof of (82) is similar to that of (70), and is omitted.

Appendix V: Proof of Theorem 3

We first derive $a_i=P(N_{1\,i+1}=1)$. Through solving the difference equation in (65), it can be found that

$$f_m=\frac{\sqrt{1+4\theta^2}+1}{2\sqrt{1+4\theta^2}}\left(\frac{1+2\theta+\sqrt{1+4\theta^2}}{2}\right)^{m-1} + \frac{\sqrt{1+4\theta^2}-1}{2\sqrt{1+4\theta^2}}\left(\frac{1+2\theta-\sqrt{1+4\theta^2}}{2}\right)^{m-1}. \quad (83)$$

Let $\mathbf{T}_{(l),ring}$ be the set of all traffic patterns on a ring network with l nodes and a route set that includes all possible link-disjoint shortest paths between each pair of nodes in the network. Let $g_m=\sum_{\mathbf{T}_{(m),ring}} \theta^{sum(\mathbf{W}_{(m),ring})}$ By counting different ways of using one single link in the ring network, we have

$$g_k=f_k+\theta f_k+2\theta^2 f_{k-1}+\ldots+(k-1)\theta^{k-1}f_2. \quad (84)$$

Therefore, $a_i=P(N_{A_1A_{i+1}}=1)=\theta^i f_{k-i+1}/g_k$. Using the lower and upper bound of $M_{f_{sd}}$ for the ring network in (25), we obtain Theorem 3.

Appendix VI: Proof of Theorem 4

From (75)-(78),

$P(N_{1\,m}=1)=\theta t_{k-1}/t_k;\ P(N_{A_1A_2}=1)=\theta^2 t_{m-2}/t_m;\ \forall k>3$.

Using the lower and upper bound of $M_{f_{sd}}$ for the star network in (27) and (28). We can obtain Theorem 4.

Appendix VII: Proof of Theorem 5

Since $M=\sum_{f_{sd}} M_{f_{sd}}\ P(N_{sd}=1)/2\,|\,\mathbf{R}\,|$, we have

$$M\le \frac{\sum_{f_{sd}} P(N_{sd}=1)}{2\,|\,\mathbf{R}\,|}\max_{f_{sd}}\{M_{f_{sd}}\}. \quad (85)$$

Let $\mathbf{N}=(N_{sd}: r_{sd}\in\mathbf{R})$, and $E[\,]$ stands for expectation. Since each connection consists of two flows, then

$$\sum_{f_{sd}} P(N_{sd}=1)=2sum(\mathbf{N}).$$

Thus, to prove Theorem 5, it suffices to show that

$$E(sum(\mathbf{N}))\le \rho\,|\,\mathbf{E}\,|.$$

Clearly, we have

$$E[sum(\mathbf{N})]=\sum_{\mathbf{W}} E[sum(\mathbf{N})\,|\,\mathbf{W}]P(\mathbf{W}),$$

where $\mathbf{W}=(W_{ij}, i\sim j)$.

Since $E[sum(\mathbf{N})\,|\,\mathbf{W}]\le E[sum(\mathbf{W})\,|\,\mathbf{W}]$,

$$E(sum(\mathbf{N}))\le E(sum(\mathbf{W})). \quad (86)$$

As $E[sum(\mathbf{W})]=\rho\,|\,\mathbf{E}\,|$, from (86),

$$E(sum(\mathbf{N}))\le \rho\,|\,\mathbf{E}\,|.$$

It follows that

$$M\le \frac{1}{|\,\mathbf{R}\,|}\max_{f_{sd}}\{M_{f_{sd}}\}\rho\,|\,\mathbf{E}\,|.$$