



# Congestion aware ant colony optimisation algorithm for routing and spectrum assignment in flexi-grid optical burst switching networks

Joshua Oladipo<sup>1</sup> · Mathys C. du Plessis<sup>1</sup> · Tim Gibbon<sup>2</sup>

Received: 4 April 2021 / Accepted: 1 March 2023 / Published online: 28 March 2023  
© The Author(s) 2023

## Abstract

This paper presents an improved distributed ant colony optimisation algorithm for routing and spectrum assignment (RSA) on optical burst switched (OBS), flexible spectrum networks, with the spectrum continuity constraint imposed. The improved algorithm incorporates a distributed method for monitoring the congestion along the links and using this information to determine what route-spectrum combination should be selected to minimise the burst loss probability (BLP). Evaluations were performed on an optical burst switching simulator to determine the effect on BLP when a dynamic route congestion measure is considered as opposed to the use of the static route length measure. The evaluations, considering the effects of optical impairments, were made on different network topologies, with multiple spectrum widths provisioned on the network, and with different loads. The incorporation of congestion measures into the algorithm showed improvements in BLP over previous work in the range of 2–32% depending on the evaluation conditions.

**Keywords** Flex-spectrum · Optical burst switching · Ant colony optimisation · Routing and spectrum assignment

## 1 Introduction

The potential bandwidth that could be accessed over optical fibre vastly outstrips that available over any other communication medium currently in use [1, 2]. However, due to various limitations of current technology, the full potential of optical fibres is not currently being taken advantage of. On the other hand, demand for network bandwidth is steadily and quickly increasing due to the widespread adoption of various high bandwidth applications by users, and the increase in Internet access over fixed and mobile channels

[3]. With the advent and increasingly widespread use of optical technologies such as Reconfigurable Optical Add Drop Multiplexers (ROADM) and Optical Cross Connects (OXC) which allow for reconfigurable optical networks, researchers are presented with the opportunity to develop advanced control plane algorithms to derive maximum benefit from optical networks [3, 4].

Towards that end, optical burst switching (OBS) was proposed as a method of increasing the throughput of optical networks [5, 6]. This combined with flex-spectrum transmission promises significant improvements in the efficiency of modern optical fibre networks. A major issue that needs to be solved in OBS networks, towards making its adoption practical, is the routing and spectrum assignment (RSA) problem [7, 8]. RSA deals with the routing and assignment of spectrum in order to ensure that the current request is fulfilled, the burst loss probability (BLP) of future requests is minimised and ensure that available network resources are used in an efficient manner [9]. The RSA problem consists of two parts: computing a route from source to destination and then assigning a contiguous segment of spectrum on the selected route [9].

Ant colony optimisation (ACO) is a swarm intelligence algorithm which has been successfully applied to routing on

---

✉ Joshua Oladipo  
Joshua.Oladipo@mandela.ac.za  
Mathys C. du Plessis  
MC.duPlessis@mandela.ac.za  
Tim Gibbon  
Tim.Gibbon@mandela.ac.za

<sup>1</sup> Department of Computing Sciences, Nelson Mandela University, P.O. Box 77000, Port Elizabeth 6001, South Africa

<sup>2</sup> Centre for Broadband Communication, Nelson Mandela University, Port Elizabeth, South Africa

networks by researchers [9, 10]. It allows for dynamic routing over a network while being able to adapt to changes on the network. In this paper, the flexible spectrum ant colony (FSAC) algorithm of Gravett, Du Plessis, and Gibbon [9] is extended by introducing congestion measures into the route and spectrum selection process in order to make the algorithm congestion aware and further minimise the achieved BLP. This study is performed using the simulator presented in [11]. The simulator was built to simulate Dense Wavelength Division Multiplexing (DWDM) and flexi-grid OBS networks in the presence of optical impairments and was validated against a reduced link load model for OBS networks. This is also the first evaluation of FSAC on an analytically validated simulator.

The rest of this paper is structured as follows: Section 2 gives a brief introduction to the various related topics, Sect. 3 presents the methods used to introduce congestion information into FSAC, Sect. 4 presents the experimental procedure used to compare the algorithms, and Sect. 5 presents the results of the performance comparison. Finally, conclusions are given in Sect. 7.

## 2 Related work

In this section, the technologies relevant to this study are introduced with particular focus on their motivations and challenges (Sects. 2.1 and 2.2). After which, previous work presented in [9] on the application of ACO to the RSA problem on flex-spectrum OBS networks is also introduced (Sect. 2.4).

### 2.1 Transparent flexible spectrum optical networks

WDM is a method of increasing the transmission capacity of an optical fibre, by allowing the parallel transmission of multiple streams of data with potentially different transmission speeds over different spectrum segments (wavelengths) across an optical fibre [12]. It has been standardised into three schemes by the International Telecommunication Union; two fixed grid schemes known as Coarse Wavelength Division Multiplexing (CWDM) and Dense Wavelength Division Multiplexing (DWDM), and a flexi-grid scheme [13]. DWDM was created to supersede CWDM due to improved technology, replacing the 18 widely spaced channels of CWDM with over a 100 channels spaced at 50 GHz or 100 GHz. Flexi-grid, in turn, was created to supersede DWDM, by allowing the use of arbitrarily positioned and sized channels, made up of smaller concatenated channels (whose size are multiples of 12.5 GHz). This has the advantage of reducing wastefulness of spectrum due to large guard bands (unused spectrum present in fixed grid slots to protect channels from interfering with each other; this is

known as inter-channel crosstalk), and efficient deployment of increased transmission rates and multi-line rate networks. However, due to the lack of large guard bands as are present in fixed grid systems, the impact of impairments (particularly inter-channel crosstalk) which might prevent signals from reaching their destination is more prominent than in flexi-grid systems [11, 14, 15]. It was shown in simulation that flexi-grid transmissions are affected by impairments more than fixed grid transmissions [11].

WDM networks can be classified as opaque, transparent or translucent [16]. In opaque networks, optical–electrical–optical (OEO) conversion occurs at each node; all incoming optical transmissions are converted to electrical signals before being retransmitted across another link if they have not arrived at their destination. This allows for spectrum conversion at intermediate nodes (i.e. switching a transmission from its current channel to a free one, if the required spectrum segment is being occupied by another transmission) and retransmission at intermediate nodes, hence improving the reach of transmissions. On the other hand, in transparent networks (also called all-optical networks), intermediate nodes on the way to the destination can be optically bypassed by optical channels which have not arrived at their destination, while transmissions terminating at that node are converted to electrical signals. Thus, a transmission between two nodes can traverse the network without undergoing OEO transmission. Translucent networks are simply WDM networks which have a mix of transparent and opaque nodes.

Transparent networks are more scalable and reliable than opaque networks due to large reduction in required electronics (particularly transponders) yielding reduced cost, space, heat dissipation and power requirements [3, 16]. They are also much easier to extend, as extending a node only requires the addition of more optical transponders [3, 16]. Finally, and most importantly, they are agnostic to the spectrum division, modulation scheme, line rate and protocol being used; that is in an opaque network every node needs to support all transmission line rates and modulation format [3, 16]. However, in a transparent optical network, only the source and destination nodes need to support the formats being used for transmission. Disadvantages of transparent networks are: transmissions are not regenerated due to absence of OEO conversion at each node, hence increasing the effects of impairments on transmissions (potentially reducing the reach of transmissions) and reducing the performance monitoring ability of the network [16, 17]. An incoming transmission into a network would tie up a segment of spectrum across multiple links, preventing it from being used by subsequent transmissions which arrive within the duration of the first transmission and have intersecting paths on the network. This is called the spectrum continuity (SC) constraint [16, 17].

The SC constraint has the implication that transmissions might be lost even though there is enough free spectrum available on the required links; due to the necessary spectrum segment being fully or partly occupied along the route of the transmission. In flexi-grid networks, the SC constraint also leads to fragmentation, where due to the mixture of transmissions with various transmission speeds on the network (faster transmissions require more spectrum), free spectrum along a route can become isolated, hence preventing the assignment of such spectrum to new transmissions due to the isolated segments not being large enough to service new requests [14, 17].

### 2.2 Optical burst switching

Current optical networks use a switching paradigm known as optical circuit switching (OCS), where all the bandwidth on a channel (where a channel could be a fibre, a spectrum slot, or a time slot in a time division multiplexed (TDM) system) is fully dedicated to transmission between a pair of nodes [1, 5]. The channels may be assigned statically or as required for a given amount of time. The disadvantage of this is that, for bursty (irregular) data, the bandwidth of a channel is reserved for a given node, irrespective of whether the node has data to transmit or not. A node might not need to send data, or might only need to send data across a channel for a short period of time, leaving the remaining channel slot unused, while other applications might currently need the bandwidth; this results in inefficiency. A more efficient way would be to use optical packet switching (OPS), where packets sent over a given spectrum slice are independently switched towards their destination (this is known as statistical multiplexing) [18, 19]. However, true optical packet switching is not currently possible due to the absence of cost effective optical buffers which would allow the buffering of packets to enable contention resolution at core nodes, as well as efficient methods for processing each packet within the optical domain [18, 19].

Optical burst switching (OBS) could be considered as a compromise to improve the efficiency of optical networks using currently available technology [6, 18, 20, 21]. In an OBS network, packets are buffered and assembled into bursts, at network edge nodes, according to various parameters. When a burst is ready to be sent, the node sends a packet, known as a Burst Control Packet (BCP) on a reserved channel called the control channel, in order to reserve the required bandwidth and setup the switching at intermediate core nodes. The burst is then sent after a certain offset time, in order to give the BCP time to finish setting up the switching at the intermediate nodes. The reserved bandwidth for a transmission is released after the burst has traversed the network, in order to allow its use by another burst. OBS leads to more efficient usage of network resources, since bandwidth

on the network is not tied up at any node for a long period, like in OCS [18]. OBS has been shown to offer higher data throughput and lower blocking rates compared to OCS in simulation studies performed by [22] and [23]

Burst loss in an OBS network can occur due to two factors, spectrum contention between two or more burst reservations on a route (leading to the loss of at least one burst), and linear and nonlinear optical impairments that could render the data sent in a burst unreadable at the destination node [24, 25]. Solving the RSA problem in a dynamic, efficient manner is one of the main challenges for OBS on flexi-grid networks [15], and is known to be NP-hard. The performance of RSA algorithms is typically measured using the burst loss probability (BLP) experienced by a network; this is defined as the ratio of number of burst lost on the network to the number of burst sent on the network.

### 2.3 Flexi\_obsmodules simulator model

The Flexi\_Obsmodules simulator [11] was built to simulate the real-time performance of DWDM and flexi-grid OBS networks in the presence of optical impairments. It is built using the Omnet++ framework [26] and designed to be easily extensible, hence, allowing for the rapid design and implementation of algorithms for various aspects of OBS networks. It has been validated against a reduced link load model to ensure credibility of the results obtained.

The simulator consist of three major components which encapsulate the various functions of an OBS network as presented in Sect. 2.2; namely the edge node, core node and fibre modules. The edge node encapsulates the burst assembly and routing functions; the core node encapsulates the network switching functions and could also perform routing functions, in the case of a distributed routing algorithm; finally, the fibre module encapsulates the details of sending data over a fibre link.

The simulator models the influence of linear and nonlinear optical impairments on transmissions over a length of fibre using Eq. (1). Equation (1) is the power penalty model presented in [27]; it gives the loss of power (in dB) experienced by a transmitted signal across a length of fibre due to linear and nonlinear impairments

$$Penalty(dB) = AL + kL \sum_{i \in T_f \setminus \{s\}} \frac{b_s 10^{\frac{P_i}{10}}}{b_i 10^{\frac{P_s}{10}} (f_i - f_s)} \tag{1}$$

where constant  $k = 4.78$  (Assuming On-Off keying modulation),  $A$  is the attenuation constant in dB/km,  $L$  is the fibre length in km,  $T_f$  is the set of signals simultaneously traversing the fibre between the start and end time of the signal,  $b_s$  is the bit rate of the signal,  $b_i$  is the bit rate of the signal causing the interference,  $P_i$  is the power of the signal causing the

interference in dBm,  $P_s$  is the power of the signal in dBm,  $(f_i - f_s)$  is the difference between the central frequencies of the interfering signal and the signal in GHz.

## 2.4 Flexible spectrum ant colony algorithm (FSAC)

Ant colony optimisation is a swarm intelligence technique which is based on the foraging behaviour of ants. It seeks to mimic the emergent, optimal path-finding foraging behaviour which can be observed in ant colonies, due to the stigmergic communication and collaboration of individuals within the colony [28, 29]. It is one of multiple approaches of interest in flex-spectrum optical networks particularly with OBS, as a way of solving the RSA problem in such networks [30]. In nature, ants accomplish the task of foraging for food by depositing small amounts of pheromones on their path, when returning to their nest from a food source. When an ant discovers a food source, it carries some food to the nest while depositing pheromones along its path. Other ants that come across the path might then probabilistically select the path based on the strength of its pheromone concentration. The more ants select a given path, the stronger its pheromone concentration will be and the greater the probability of other ants selecting it becomes; while the pheromone concentration of less used paths become weaker over time as the pheromones deposited on them evaporate. A factor that influences the pheromone concentration of a path is the length of the path; this is due to the fact that ants will return quicker along a shorter path than along a longer one. Hence, reinforcing the pheromones along that path faster and making the pheromone concentration higher than other paths. In this way, the shortest path is chosen by the ants.

FSAC is an OBS-based RSA algorithm for OBS networks, specifically designed to effectively route data over flex-spectrum networks under the SC constraint, while considering the effects of network impairments [9]. The algorithm is evaluated in both fixed grid and flex-spectrum optical networks and is found to out-perform both shortest path routing (SPR) and ant colony routing and wavelength assignment (ACRWA) algorithms [10] on all scenarios investigated.

Each edge node in a network which uses FSAC holds a pheromone table  $\Xi$ , which contains  $n$  unique entries for each possible destination (where an entry is identified by its route, spectrum pair). Each entry  $\varepsilon_i$  is a tuple  $\langle \mathcal{N}_{nmi}, \lambda_i, \kappa_i, \chi_i, \tau_i, \eta_{nmi} \rangle$  consisting of a route  $\mathcal{N}_{nmi}$  between the current edge node  $n$  and a destination  $m$ , a central frequency  $\lambda_i$ , a success counter  $\kappa_i$ , a fail counter  $\chi_i$ , a pheromone value  $\tau_i$  and the desirability value of an entry  $\eta_{nmi}$ .  $\eta_{nmi}$  is given as the reciprocal of the length of the route  $\mathcal{N}_{nmi}$ . The entries are created by randomly selecting a route from a candidate list of  $K$ -shortest paths, randomly selecting a spectrum slot from the set of available spectrum,

initialising all counters to zero and initialising the pheromone values to a random value between 0 and 1.

When a burst request is made, the source node  $s$  determines which entry to use for the transmission by generating a random value  $r \sim U(0, 1)$ , which determines if the burst is going to exploit previous entries or explore a new one as follows

- If  $r < \alpha_1$ , then an entry  $\hat{\varepsilon}$  is chosen according to the following expression

$$\hat{\varepsilon} = \operatorname{argmax}_{\varepsilon_j \in \Xi_{nm}} \{ \tau_j \eta_j^\beta \} \quad (2)$$

where  $\alpha_1 \in (0, 1]$  is the probability that the algorithm will exploit the best entry according to Eq. (2) and  $\beta$  is a constant.

- If  $r < (\alpha_1 + \alpha_2)$ , then the probability of an entry  $\varepsilon_i$  being selected is given by

$$\rho_i = \frac{\tau_i \eta_i^\beta}{\sum_{\varepsilon_j \in \Xi_{nm}} \tau_j \eta_j^\beta} \quad (3)$$

where  $\alpha_2 \in (0, 1]$  is the probability that the algorithm will explore new entries using roulette wheel selection as shown in 3.

- If  $r \geq (\alpha_1 + \alpha_2)$ , then a new entry is generated by randomly selecting a central frequency and a path (with source  $s$  and destination  $d$ ) from the candidate list of  $K$ -shortest paths. The new entry is then used to route the current request, and replace the entry for routing between  $n$  and  $m$  with the lowest pheromone value in the pheromone table.

A BCP is sent along the chosen route and spectrum slice, in order to set up the switching configuration at each intermediate node. The burst is sent after the offset time from the BCP has elapsed. After a burst has been successfully delivered, feedback is sent in the form of an acknowledgement message (hereby referred to as BCP-ACK) to inform the source node of a successful or failed delivery, to allow it update the pheromone concentration of an entry  $m$ . The authors of [9] tested various equations for calculating the new pheromone value. It was found that, of the equations that were tested, the following equation best minimised the BLP:

$$\tau_i = e^{\psi \frac{\kappa_i + 1}{\kappa_i + \chi_i + 1}} \quad (4)$$

where  $\psi$  is a preset constant and  $\kappa_i$  and  $\chi_i$  are the count of the successful and failed transfers at time  $t$  made using the entry  $\varepsilon_i$ .

After an intermediate node forwards the BCP it starts a time-out timer, which is only stopped after a BCP traversal acknowledgement is received from the receiving node. The time-out timer is set to some multiple of the link traversal time. If no acknowledgement is received by the time the

timer runs out, the resources reserved for the burst are released and a BCP-ACK informing the source node of a failure is sent to the source node, along intermediate nodes, allowing them to also free up reserved resources. This is to deal with the situation where a BCP is lost due to optical impairments.

### 3 Incorporating congestion information

The desirability of a route in FSAC is given as the reciprocal of the route length. This means that the algorithm will tend to favour shorter routes. However, the desired effect the authors would like to observe is a load balancing effect, where nodes seek to efficiently exploit underutilised resources (routes and spectrum). To that end, congestion information of routes is introduced into the algorithm. This is a dynamic real-time estimation of the network state as opposed to the static measure previously used. Using a dynamic measure should make the algorithm more sensitive to the state of the network.

Congestion information is recorded by BCP-ACK’s of successful bursts travelling along the reverse route to the source node. On arriving at the source node  $n$ , the collected information is used to estimate the congestion along a route. The estimated congestion value serves as the desirability value for the route, and all entries which use the route followed by the corresponding burst, are updated with the desirability value. In this section, three novel schemes for estimating the congestion of a route based on the reduced link load model discussed in [11] are presented; followed by the description of an algorithm with no access to network information which will serve as a control.

#### 3.1 CM-FSAC operation

Three methods of congestion measurement (denoted with the prefix “CM”) are proposed. Two of them estimate the congestion along a route, using the amount of free spectrum along the route, while the other estimates the congestion along a route using the number of BCP’s lost along a route due to contention.

- *CM1*—Measures the congestion along a route from node  $n$  to node  $m$  as

$$\eta_{nmi} = B_{nm} = 1 - \prod_{k=1}^J (1 - B_{nmk})$$

with

$$B_{nmk}(M_k) = \frac{M_k}{M_k + 1} \tag{5}$$

where

$$M_k = \frac{W - U_k}{W}$$

Equation (5) is an adaptation of the Erlang-B formula for analytically calculating the burst loss over a network which maintains the SC constraint. Its derivation is presented in [11]. The congestion over a link is estimated as  $B_{nmk}$ , a function of  $M_k$  (the fraction of bandwidth occupied on link  $k$ ).  $W$  is the spectrum width and  $U_k$  is the available spectrum on fibre link  $k$ .  $J$  is the number of links on the path.

- *CM2*—Measures the congestion along a route from node  $n$  to node  $m$  using the same procedure as *CM1*, except that  $M_k$  becomes the ratio of burst lost on a link due to contention, measured by calculating the percentage of failed BCP’s. That is,

$$M_k = \frac{\text{Count of failed reservations on link } k}{\text{Count of reservations for link } k} \tag{6}$$

It is assumed that control channels are impairment free; hence, BCP’s cannot be lost for any other reason than contention. Burst lost due to impairments are not considered in this measure, due to the fact that the majority of lost burst in a network are still due to contention as shown in [11]. And the impairments experienced by transmissions on a route are considered to be a function of the number of transmissions being made on that route (as can be seen in Eq. 1). Hence, using only the percentage of BCP’s lost due to contention may be sufficient to estimate how busy a route is.

- *CM3*—*CM3* uses the percentage of BCP’s lost along a route. It measures this by collecting the number of BCP’s lost at links along the route due to contention, and total number of BCP’s sent along the route, and calculating the percentage of BCP’s lost along that route. That is,

$$\eta_{nmi} = \frac{\text{Count of failed reservations on } \mathcal{N}_{nmi}}{\text{Count of reservations on } \mathcal{N}_{nmi}}$$

Only BCP’s lost due to contention are considered, for the same reasons given for *CM2*.

#### 3.2 Information deprived FSAC

In order to obtain a baseline on the value of network information to the performance of the algorithm, FSAC and *CM-FSAC* are compared against an algorithm with no access to network information, such as route lengths or route congestion information. This algorithm is denoted as information

deprived FSAC (ID-FSAC). The tuple used by FSAC is redefined as  $\langle \mathcal{N}_{nm}, \lambda_i, \kappa_i, \chi_i, \tau_i, \Gamma_i \rangle$ ; where  $\Gamma_i$  is the number of times entry  $i$  has been used. Equation (2) is replaced with

$$\hat{\varepsilon} = \operatorname{argmax}_{\varepsilon_j \in \Xi_{nm}} \{\tau_j\} \quad (7)$$

and Eq. (3) is replaced with

$$\rho_i = \frac{\sum_{\varepsilon_j \in \Xi_{nm}} \Gamma_j - \Gamma_i}{\sum_{\varepsilon_j \in \Xi_{nm}} \Gamma_j} \quad (8)$$

The intended effect of Eq. (7) is to select an entry based only on the pheromone concentration; that is the entry with the highest pheromone concentration is selected. The intended effect of Eq. (8) is to probabilistically explore least used entries.

## 4 Experimental procedure

In this section, the experimental procedure used in this study is presented; starting with the network scenarios on which algorithm comparisons were performed and details on how the performance comparison was performed.

### 4.1 Network scenarios

Simulations were performed on two network topologies; the 6-node SIMPLE topology shown in Fig. 1a and the 14 node NSFNET topology shown in Fig. 1b. The control channels were made impairment free. The assumption being made is that a network operator could provide an impairment free control channel (such as in the form of a separate fibre). All simulations are performed with impairments enabled for data channels and the SC constraint imposed. The scheduling scheme used was the just-in-time (JIT) protocol [31]. In JIT, intermediate nodes perform the bandwidth reservation and switching immediately after they receive the BCP. The bandwidth is reserved for the whole period of the transmission of the burst, and any request for a reserved spectrum is

discarded. The offset time between the BCP and the burst is calculated in a route length dependent manner, as presented in [32].

For the purpose of this study, three load levels are arbitrarily defined; high, medium and low load. These are measured using the mean burst loss probability (BLP) obtained by the shortest path routing algorithm (SPR) with random spectrum selection. In SPR, the shortest path and a random free slice of spectrum are selected and attempted at the source node. High load scenarios are defined as network scenarios where SPR achieves a BLP greater than 30%, medium load as a BLP greater than 10% but less than or equal to 30%; and finally, low load as a BLP less than or equal to 10%.

In order to simulate the multi-line rate nature of flexi-grid networks, each burst is sent on a channel with a randomly assigned transmission speed, such that a specified percentage of all burst sent over the network traverse the network at a given speed. The percentage of burst that were transmitted at a given transmission speed are:

- 40% of traffic transmitted at 10 Gbps
- 30% of traffic transmitted at 20 Gbps
- 20% of traffic transmitted at 30 Gbps
- 10% of traffic transmitted at 40 Gbps.

The assumption being made by the distribution of traffic given above is that in a commercial network, fewer users require and pay for higher transmission speeds. The transmission speeds of all control channels were set to 40 Gbps. The amount of spectrum used for each line rate is determined as specified in [11].

The effect of varying the available spectrum width on the fibres was also investigated as this was found to be a factor affecting the magnitude of the improvement that can be achieved by the ACO algorithms. Simulations were run on the NSFNET and SIMPLE topologies with increasing spectrum widths, while still maintaining the same load level. The NSFNET topology is run with 200 GHz ( $16 \times 12.5$  GHz flexi-grid channels), 300 GHz ( $24 \times 12.5$  GHz flexi-grid channels) and 400 GHz ( $32 \times 12.5$  GHz flexi-grid channels)

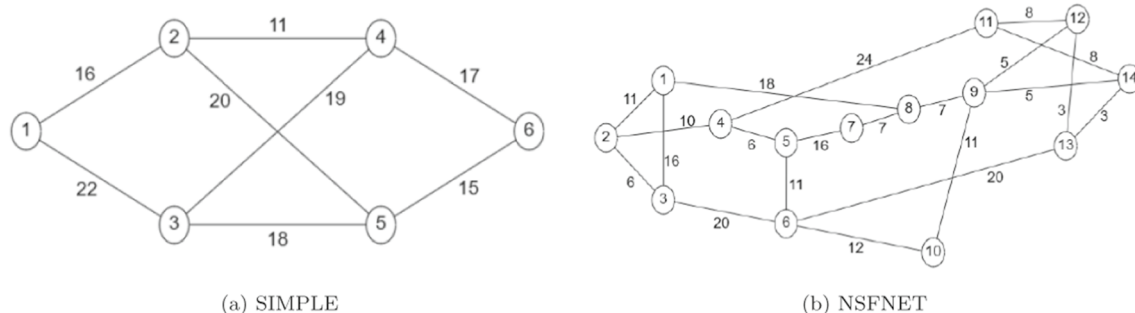


Fig. 1 Network topologies used in evaluations

spectrum widths; and the SIMPLE topology is run with 100 GHz (8 × 12.5 GHz flexi-grid channels), 200 GHz (16 × 12.5 GHz flexi-grid channels) and 300 GHz (24 × 12.5 GHz flexi-grid channels) spectrum widths.

The load level achieved on a network is dependent on the spectrum width available on the network and the packet inter-arrival time. Packets were generated using exponentially distributed inter-arrival times with the mean of the distribution set as specified in Table 1 (increasing the mean inter-arrival time decreases the load). The size of the generated packets was exponentially distributed with the mean set at 1.5 kB; 1.5 kB was chosen as this is the maximum packet size of Ethernet packets as defined by the IEEE 802.3 standard. The burst assembly scheme used was size based, with the maximum burst size being 15 kB.

### 4.2 Performance comparison

Performance evaluations were performed to determine the performance of the various algorithms on the scenarios presented above, using the parameters given in Table 4. The parameter values used in the performance comparison were determined by performing a grid search [33]. Each configuration (a topology, a load, a spectrum width and an algorithm) was run 30 times with different seeds, in order to ensure the reliability of the results. The three spectrum widths investigated on each network are labelled with the letters A, B, C in order to keep the table concise. For the simple topology, the spectrum widths are 100 GHz (A), 200 GHz (B) and 300 GHz (C). For the SIMPLE topology, the spectrum widths are 100 GHz (A), 200 GHz (B) and 300 GHz (C). For the NSFNET topology, the spectrum widths are 200 GHz (A), 300 GHz (B) and 400 GHz (C).

The values for  $\alpha_1$  and  $\alpha_2$  were set to  $\alpha_1 = 0.98$  and  $\alpha_2 = 0.0175$  for all the algorithms in order to satisfy Eq. (9). During the parameter tuning, it was found that  $\alpha_2$  should satisfy Eq. (9)

$$\alpha_2 > 1 - (\alpha_1 + \alpha_2) \tag{9}$$

**Table 1** Mean inter-arrival time values used to achieve various loads in performance comparisons

Topology	Spectrum width (GHz)	High (ms)	Medium (ms)	Low (ms)
SIMPLE	100 (A)	0.065	0.11	0.7
	200 (B)	0.033	0.055	0.35
	300 (C)	0.023	0.038	0.25
NSFNET	200 (A)	0.09	0.2	0.98
	300 (B)	0.062	0.134	0.65
	400 (C)	0.048	0.105	0.5

Equation (9) says that the probability of exploring entries within the entry table should be higher than the probability of creating new entries. Hence, it was observed that if the probability of creating new entries is greater than or equal to the probability of exploring entries, the algorithm will not have enough opportunities to explore newly created entries; hence, the performance of the algorithm will be negatively affected. Only the number of entries is given for ID-FSAC as it was observed that only variations in the number of entries affected its performance.

The results obtained from the simulations are the BLP values over time for all experiments that were performed. The final BLPs for each algorithm were then averaged. The best performing algorithm is the algorithm with the lowest mean BLP. The algorithms were also subjected to the Mann–Whitney U test against each other, in order to ensure that there is a statistically significant difference in their performance [34].

## 5 Results

Tables 2 and 3 present the mean BLPs for all scenarios for the SIMPLE and NSFNET topology, respectively. The best performing algorithm is indicated for each scenario is indicated in bold. The BLPs for the SIMPLE topology were obtained at 800 s, while for the NSFNET topology they were obtained at 430 s. These values were found to be the minimum time required to ensure that a steady state is reached by each scenario. For all scenarios, one of the “CM” algorithms was found to be the best performing ACO algorithm. CM1 was found to have the lowest mean BLP in the majority of scenarios, making it the best overall algorithm. Using the Mann–Whitney U test, it was found that there is a statistically significant difference in the performance of the algorithms.

CM1 performs best overall in three of the nine scenarios on the SIMPLE network, and in six of the nine scenarios on the NSFNET network. CM2 performs best overall in two of the nine scenarios on the SIMPLE network, and zero of the nine scenarios on the NSFNET network. CM3 performs best on three of the nine scenarios on the SIMPLE network, and zero of the nine scenarios on the NSFNET network. SPR was the best performer on one of the nine scenarios on the SIMPLE network and on three of the nine scenarios on the NSFNET network. FSAC was not the best performer for any of the scenarios. The performance of the ACO algorithms (except ID-FSAC) is seen to be comparable, as the difference between FSAC’s BLP and that of the best performing CM algorithms is on average 0.01 (that is, a 1% reduction in number of burst lost). The performance of FSAC is also found to be similar to that reported in [9].

**Table 2** Mean BLPs obtained from performance comparisons on the SIMPLE topology

Spectrum width	Load	SPR	FSAC	CM1	CM2	CM3	ID-FSAC
100 GHz	High	0.3550	0.3491	0.3516	0.3490	<b>0.3484</b>	0.3597
	Medium	0.2531	0.2488	0.2495	<b>0.2459</b>	0.2470	0.2623
	Low	<b>0.0530</b>	0.0593	0.0552	0.0567	0.0657	0.0719
200 GHz	High	0.3537	0.2736	0.2746	0.2731	<b>0.2724</b>	0.2812
	Medium	0.2539	0.1709	<b>0.1690</b>	0.1700	0.1705	0.1872
	Low	0.0523	0.0382	<b>0.0349</b>	0.0359	0.0438	0.052
300 GHz	High	0.3525	0.3125	0.3054	<b>0.3039</b>	0.3072	0.3175
	Medium	0.2547	0.1774	0.1700	0.1686	<b>0.1674</b>	0.1835
	Low	0.0511	0.0310	<b>0.0216</b>	0.0228	0.0293	0.0371

Best result (smallest value) for each scenario indicated in bold

**Table 3** Mean BLPs obtained from performance comparisons on the NSFNET topology

Spectrum width	Load	SPR	FSAC	CM1	CM2	CM3	ID-FSAC
200 GHz	High	0.3538	0.3472	<b>0.3466</b>	0.3473	0.3569	0.3731
	Medium	<b>0.2037</b>	0.2183	0.2040	0.2106	0.2258	0.2455
	Low	<b>0.051</b>	0.0670	0.0588	0.0628	0.0884	0.1089
300 GHz	High	0.3514	0.3308	<b>0.3219</b>	0.3234	0.3341	0.3481
	Medium	0.2056	0.2032	<b>0.1911</b>	0.1944	0.2124	0.2287
	Low	<b>0.0519</b>	0.0637	0.0547	0.0589	0.0785	0.0972
400 GHz	High	0.3496	0.3076	<b>0.2980</b>	0.2983	0.3111	0.3249
	Medium	0.2021	0.1853	<b>0.1727</b>	0.1753	0.1945	0.2104
	Low	0.0515	0.0598	<b>0.0510</b>	0.0541	0.0734	0.0899

Best result (smallest value) for each scenario indicated in bold

Table 2 presents the results of the experiments performed on the SIMPLE topology. At low load, CM1 is found to be the best performing ACO algorithm. All ACO algorithms perform worse than SPR at the 100GHz spectrum width mark. However, as the spectrum width increases the ACO algorithms perform better than SPR, with better performance at 300 GHz than at 200 GHz. At medium and high loads, CM2 and CM3 perform better than CM1, except at 200 GHz where CM1 is the best performing algorithm.

Table 3 presents the results of the experiments performed on the NSFNET topology. At low load, CM1 is found to be the best performing ACO algorithm, followed by CM2, FSAC, CM3 and ID-FSAC. This order holds for all spectrum widths. At 200 and 300 GHz, none of the ACO algorithms perform better than SPR; however, as the spectrum width increases, the performance of all the ACO algorithms improves, with CM1's performance being better than SPR at 400 GHz. At medium load, CM1 is the best performing ACO algorithm, followed by CM2, FSAC, CM3 and ID-FSAC. At 200 GHz, none of the algorithms perform better than SPR; however, as the spectrum width increases, the performance of all the ACO algorithms improves and CM1 and FSAC begin to perform better SPR. At high load, CM1 is the best performing ACO algorithm at all spectrum widths, followed by CM2, FSAC, CM3 and ID-FSAC. This order holds as the

spectrum width increases. CM1 is better than SPR for all spectrum widths, and FSAC either matches or is better than SPR for all spectrum widths (Table 4).

Figure 2 shows the performance over time of the algorithms tested for four scenarios on the SIMPLE network. These curves were obtained by taking the average performance of an algorithm over time for each scenario. From the graphs, it can be observed that the performance of SPR remains constant over time as expected (due to it not performing any intelligent actions), and the performance of the ACO algorithms improves exponentially over time due to the processes of the algorithms; this is true for all other scenarios not shown. It can also be observed that as the spectrum width and load increases, the rate of convergence also increases for all ACO algorithms.

ID-FSAC is the worst performing ACO algorithm, as expected. However, it performs better than SPR on six of the nine scenarios on the SIMPLE network, but only two of the nine scenarios on the NSFNET network. On the SIMPLE topology, ID-FSAC does not perform better for any of the 100 GHz scenarios; however, for higher spectrum width, ID-FSAC performs better than SPR. On the NSFNET topology, ID-FSAC performs better than SPR for only the high loads on the 300 and 400 GHz spectrum widths; however, the difference in performance between SPR and ID-FSAC,



**Table 4** Parameter values used for all algorithms in performance comparisons

Algorithm	Topology	Load	A			B			C		
			$\beta$	$\psi$	$N$	$\beta$	$\psi$	$N$	$\beta$	$\psi$	$N$
FSAC	SIMPLE	High	0.4	1.36	12	0.2	1.36	12	0.4	1	12
		Medium	0.6	1.36	8	0.2	1	8	0.6	1.36	12
		Low	0.2	1.36	8	0.2	1	12	0.6	1.36	12
	NSFNET	High	0.2	0.64	12	0.6	1.36	12	0.6	0.64	12
		Medium	0.4	1	12	0.4	1	12	0.8	1.36	12
		Low	1.2	0.64	12	1.2	1.36	12	0.8	1.36	12
CM1	SIMPLE	High	0.8	1.36	8	0.2	0.64	4	0.2	1.36	12
		Medium	1	0.64	8	0.2	1	12	0.2	0.64	4
		Low	1.2	1.36	8	0.8	1.36	8	1.2	0.01	12
	NSFNET	High	0.2	0.64	12	0.6	1	12	0.8	1.36	12
		Medium	1.2	1.36	12	0.8	1	12	1	1.36	12
		Low	1	0.28	12	1.2	1	12	1	1.36	12
CM2	SIMPLE	High	0.6	1	12	0.6	1.36	12	0.4	1	12
		Medium	0.4	0.64	8	0.2	1.36	4	0.2	0.64	12
		Low	0.4	0.1	12	0.6	0.1	8	1.2	0.28	12
	NSFNET	High	1.2	0.64	12	1	0.64	12	1.2	1	12
		Medium	1	0.01	12	0.4	0.1	12	1	0.28	12
		Low	0.8	0.01	12	1.2	0.01	12	0.8	0.1	12
CM3	SIMPLE	High	1	1.36	4	0.8	1.36	4	0.8	1	12
		Medium	0.2	1.36	12	0.6	1	8	1	0.64	8
		Low	1.2	0.1	8	0.8	0.28	12	0.2	0.1	12
	NSFNET	High	0.2	0.28	12	1	1	12	0.2	0.1	12
		Medium	0.4	1	12	0.8	0.28	12	1	0.64	12
		Low	1.2	0.64	12	0.2	0.1	12	0.2	0.28	12
ID-FSAC	SIMPLE	High	–	–	12	–	–	12	–	–	8
		Medium	–	–	12	–	–	12	–	–	4
		Low	–	–	12	–	–	12	–	–	12
	NSFNET	High	–	–	8	–	–	12	–	–	12
		Medium	–	–	12	–	–	12	–	–	12
		Low	–	–	12	–	–	12	–	–	12

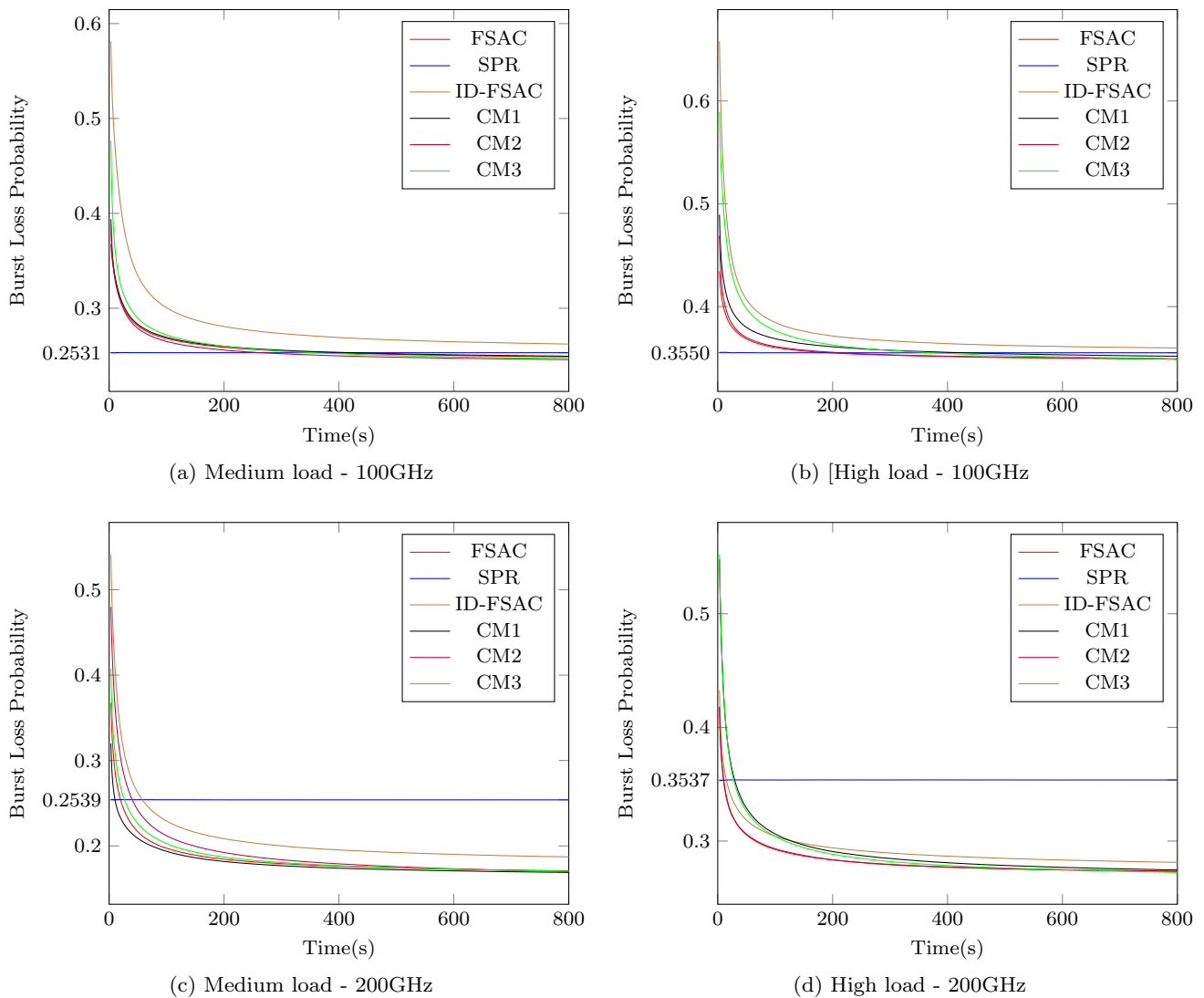
for the loads where ID-FSAC does not better SPR, becomes smaller as the spectrum width increases.

### 6 Discussion

A common trend observed on both networks is that the effectiveness of the ACO algorithms improves as the provisioned spectrum width increases. For example, at 200 GHz and high load on the NSFNET network, CM1 achieves a 2% improvement in BLP over SPR; while at 400 GHz CM1 achieves a 14% improvement in BLP over SPR. An explanation for this could be that as the spectrum width increases there are more underutilised spectrum resources on the network, which the ACO algorithms are able to take advantage of. It is also observed that the ACO algorithms seem to perform worse at low loads, as the magnitude of the performance improvements on low load

scenarios is less than that at high loads; and in some cases none of the ACO algorithms perform better such as at 100 GHz low load on the SIMPLE topology, and 200 GHz low and medium load on the NSFNET topology, and 300 GHz medium load on the NSFNET topology. A reason for this could be that less information is being obtained from the network at low loads due to less burst being transmitted (the route quality is measured by the BCP-ACKs, see Sect. 3), so the ACO algorithm’s judgement of the effectiveness of each route is less accurate. The performance of ID-FSAC compared to SPR clearly reflects the above two observations. Hence, it could be taken as an indication of the difficulty of each scenario; with high load, high provisioned spectrum width scenarios being less difficult, and low load, low provisioned spectrum widths being more difficult.

The improved performance of CM1 over CM2 and CM3 could be due to CM1 giving a more timely report of the state



**Fig. 2** Time wise performance of algorithms on SIMPLE topology with 100 GHz and 200 GHz spectrum width

of routes. The metrics collected along the links by CM2 and CM3 are based on data spanning the entire operating time of the network, while CM1 is based on data collected at a single moment in the networks life. Hence, the performance of CM2 and CM3 might be improved, by using only data collected over a shorter time window. The disadvantage of this being the introduction of a new time window length parameter to be optimised.

## 7 Conclusion

In this paper, the performance of the FSAC algorithm is improved by adapting it to consider network congestion measurement in its route and spectrum selection. Multiple methods of measuring the congestion on the network were proposed and evaluated. The results of the algorithm comparison showed that the congestion aware algorithms attained better network performance than that obtained

by FSAC which uses the route length of an entry (a static measure). The improvements observed were especially striking for results obtained from simulations run on the complex NSFNET network, and at low loads on both the SIMPLE and NSFNET networks.

The success of CM-FSAC signals a potential direction for improvement of ACO-based RSA algorithm. In future work, other pertinent measures such as spectrum fragmentation and fairness may be studied and incorporated into the algorithm to improve the adaptability of the algorithm to changes in a network's state. The effects of limiting the time window over which data for the CM2 and CM3 algorithms is collected may also be investigated.

**Acknowledgements** The authors acknowledge the Centre for High Performance Computing (CHPC), South Africa, for providing computational resources to this research project.

**Funding** Open access funding provided by Nelson Mandela University.

**Data availability** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Maier, M.: Optical Switching Networks. Cambridge University Press, Cambridge (2008). (ISBN 9780521868006)
2. Ramaswami, R., Sivarajan, K., Sasaki, G.: Optical Networks: A Practical Perspective (2010). (ISBN 9780123740922)
3. Jinno, M., Takara, H., Kozicki, B., Tsukishima, Y., Sone, Y., Matsuoka, S.: Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies. IEEE Commun. Mag. **47**(11), 66–73 (2009). <https://doi.org/10.1109/MCOM.2009.5307468>. (ISSN 01636804)
4. Sygletos, S., Tomkos, I., Leuthold, J.: Technological challenges on the road toward transparent networking. J. Opt. Netw. **7**(4), 321–350 (2008)
5. Chen, Y., Qiao, C., Xiang, Yu.: Optical burst switching: a new area in optical networking research. IEEE Netw. **18**(3), 16–23 (2004). <https://doi.org/10.1109/MNET.2004.1301018>
6. Verma, S., Chaskar, H., Ravikanth, R.: Optical burst switching: a viable solution for terabit IP backbone. IEEE Netw. **14**(6), 48–53 (2000). <https://doi.org/10.1109/65.885670>
7. Kishen Ajay Kumar, V., Suresh Reddy, K., Giri Prasad, M.N.: Review of contemporary literature on burst assembling and routing strategies in OBS networks. J. Opt. **47**(3), 324–331 (2018)
8. Shi, Z., Zengji, L., Hao, Q., Bing, Z.: A routing and wavelength assignment algorithm in optical burst switching networks. J. Opt. Commun. **26**(1), 24–31 (2005). <https://doi.org/10.1515/JOC.2005.26.1.24>
9. Gravett, A.S., du Plessis, M.C., Gibbon, T.B.: A distributed ant-based algorithm for routing and wavelength assignment in an optical burst switching flexible spectrum network with transmission impairments. Photonic Netw. Commun. **34**, 375–395 (2017)
10. Triay, J., Cervelló-Pastor, C.: An ant-based algorithm for distributed RWA in optical burst switching. In: ICTON 2009: 11th International Conference on Transparent Optical Networks, pp. 7–10 (2009). <https://doi.org/10.1109/ICTON.2009.5185278>
11. Oladipo, J., Gibbon, T.B.: Implementation and validation of an omnet++ optical burst switching simulator. In: 2017 International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN), pp. 1–6. IEEE (2017)
12. Iyer, S.: Traffic grooming with survivability and power-efficiency in software defined elastic optical networks. J. Opt. **47**(3), 351–365 (2018)
13. International Telecommunication Union - ITU-T. G.694.1 (02/2012), Spectral grids for WDM applications: DWDM frequency grid. *SERIES G.694.1*, pp. 1–16 (2012)
14. Wright, P., Lord, A., Velasco, L.: The network capacity benefits of flexgrid. In: 2013 17th International Conference on Optical Networking Design and Modeling (ONDM), pp. 7–12. IEEE (2013)
15. Gerstel, O., Jinno, M., Lord, A., Yoo, S.B.: Elastic optical networking: a new dawn for the optical layer? IEEE Commun. Mag. **50**(2), s12–s20 (2012)
16. Saleh, A.A., Simmons, J.M.: All-optical networking: evolution, benefits, challenges, and future vision. Proc. IEEE **100**(5), 1105–1117 (2012). <https://doi.org/10.1109/JPROC.2011.2182589>. (ISSN 00189219)
17. Simmons, J.M.: Optical Network Design and Planning. Springer, Berlin (2014). <https://doi.org/10.1007/978-3-319-05227-4>. (ISBN 9783319052274)
18. Maier, M., Reisslein, M.: Trends in optical switching techniques: a short survey. IEEE Netw. **22**(6), 42–47 (2008). <https://doi.org/10.1109/MNET.2008.4694173>. (ISSN 08908044)
19. Rouskas, G.N., Xu, L.: Chapter 1: optical packet switching. In: Optical WDM Networks: Past Lessons and Path Ahead, pp. 1–19. Kluwer, Massachusetts (2004)
20. Bjornstad, S., Gauger, C.M., Nord, M., Baert, E., Callegati, F., Careglio, D., Klinkowski, M., Marciniak, M., Solé-Pareta, J.: Optical burst switching and optical packet switching. In: Advanced Infrastructure for Photonic Networks—Extended Final Report of COST Action 266, pp. 115–154 (2003)
21. Battestilli, T., Perros, H., Carolina, N.: An introduction to optical burst switching. IEEE Opt. Commun. **8**, 10–15 (2003)
22. Xue, F., Yoo, S.B., Yokoyama, H., Horiuchi, Y.: Performance comparison of optical burst and circuit switched networks. In: Optical Fiber Communication Conference, 2005. Technical Digest. OFC/NFOEC, vol. 3, p. 3 (2005)
23. Liu, X., Qiao, C., Yu, X., Gong, W.: A fair packet-level performance comparison of OBS and OCS. In: Optical Fiber Communication Conference, pp. 48–50. Optical Society of America (2006)
24. Azodolmolky, S., Klinkowski, M., Marin, E., Careglio, D., Pareta, J.S., Tomkos, I.: A survey on physical layer impairments aware routing and wavelength assignment algorithms in optical networks. Comput. Netw. **53**(7), 926–944 (2009). <https://doi.org/10.1016/j.comnet.2008.11.014>. (ISSN 13891286)

25. Klinkowski, M., Pedro, J., Careglio, D., Pioro, M., Pires, J., Monteiro, P., Sole-Pareta, J.: An overview of routing methods in optical burst switching networks. *Opt. Switch. Netw.* **7**(2), 41–53 (2010)
26. OMNeT++ Discrete Event Simulator-Home. URL <https://omnetpp.org/>
27. Boiyo, D.K., Chabata, T., Kipnoo, R.E., Gamatham, R., Leitch, A., Gibbon, T.B.: Crosstalk penalties in optical fibre transport for next-generation flexible spectrum networks. In: Southern Africa Telecommunication Networks and Applications Conference (SATNAC), pp. 21–24 (2015). (ISBN 978-0-620-67151-4)
28. Dorigo, M., Di Caro, G.: Ant colony optimization: a new metaheuristic. In: Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406), vol. 2, pp. 1470–1477 (1999). <https://doi.org/10.1109/CEC.1999.782657>. (ISBN 0-7803-5536-9)
29. Blum, C.: Ant colony optimization: introduction and recent trends. *Phys. Life Rev.* **2**(4), 353–373 (2005)
30. Donato, E., Junior, J.C., Vieira, A., Patel, A.: A proposal of dynamic RWA using ant colony in optical burst switched networks. In: The Proceedings of the Eleventh International Conference on Networks (ICN 2012), pp. 246–252 (2012)
31. Kirci, P., Zaim, A.H.: Just-in-time, just-enough-time and horizon signalling protocols on optical burst switches. *Opt. Appl.* **36**(1) (2006). (ISSN 00785466)
32. Teng, J., Rouskas, G.N.: A comparison of the JIT, JET, and horizon wavelength reservation schemes on a single OBS node. In: Proceedings of the First International Workshop on Optical Burst Switching (2003)
33. Elgeldawi, E., Sayed, A., Galal, A.R., Zaki, A.M.: Hyperparameter tuning for machine learning algorithms used for Arabic sentiment analysis. Informatics, Multidisciplinary Digital Publishing Institute (2021)
34. Mann, H.B., Whitney, D.R.: On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **18**, 50–60 (1947)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Joshua Oladipo** is a Computer science PhD student at Nelson Mandela University. He obtained his MSc Computer science in 2018, BSc (honours) Computer science in 2015. His current field of study is the application of evolutionary computing algorithms to routing over optical networks.



**M.C. du Plessis** is an academic in the Department of Computing Sciences at Nelson Mandela University. He obtained a PhD from the University of Pretoria in 2012. His research interests include Ant Colony Optimization, Neural Networks, Evolutionary Robotics and Differential evolution.



**Tim Gibbon** is Director of the CBC and Associate Professor in the Department of Physics. In 2007, he received a PhD in Physics from NMMU in the field of optical communication. From 2008–2010, he was Postdoctoral Researcher at DTU Fotonik in Denmark, working closely with industry developing next-generation WDM-PON technologies. Tim is an NRF-rated scientist and has published over 70 papers on topics including

long-haul impairments, metro-access technologies, passive optical networks, FTTH, photonic UWB, and optical fibre telescope networks. He currently leads Nelson Mandela University's work on the SKA Project, with further industry and academic collaborations across Africa, Europe, China and Australia.