Yarr, N. and Ceriotti, M. (2018) Optimization of inter-satellite routing for real-time data download. IEEE Transactions on Aerospace and Electronic Systems, 54(5), pp. 2356-2369.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Deposited on: 16 March 2018

Optimization of Inter-Satellite Routing for Real-Time Data Download

Neil Yarr and Matteo Ceriotti

Abstract—The objective of this study is to develop a strategy to maximise the available bandwidth to Earth of a satellite constellation through inter-satellite links. Optimal signal routing is achieved by mimicking the way in which ant colonies locate food sources, where the 'ants' are explorative data packets aiming to find a near-optimal route to Earth. Demonstrating the method on a case-study of a space weather monitoring constellation; we show the real-time downloadable rate to Earth.

Index Terms — Ant Colony Optimization, Constellation, CubeSat Mission, Inter-Satellite Routing, Optimization.

I. INTRODUCTION

INTER-satellite links (ISLs) are telecommunication pathways between different satellites which allow a group or formation of satellites to effectively become a network of relay nodes. ISLs can be used to share data amongst different nodes of a network; one possible aim is to increase and possibly maximize the bandwidth between a ground station and a spacecraft in the network, not necessarily in direct sight of the ground station, at any given time. Individual satellites do not need to wait for the brief window of time in which a particular ground station is within the field of view of its antenna, but can transfer data to other satellites at different positions in the constellation, either on the orbit (intra-orbital) or on different orbits (inter-orbital). With these links in place, satellites in large-enough constellations can communicate with relevant ground stations in quasi-real-time, regardless of where they are in relation to ground stations. It is clear that the extent of the usefulness of ISLs depends on the effectiveness of the routing strategy employed. The main difficulty in utilizing ISLs as a networking solution is the fact that in most satellite constellations, the network topology is time-varying; links will constantly be found/lost as each satellite progresses along its own orbit, hence the effectiveness of the routing strategy becomes key to exploiting the availability of ISLs.

With the increasing popularity of CubeSats as a low-cost, standardized and relatively quick to manufacture solution to space access, the question of inter-satellite telecommunications must be revisited and applied to this type of platform. Their small size and power available imply that CubeSats have much lower communication capabilities than traditional satellites; limitation in attitude control and size means that only low-gain,

omnidirectional antennas can be used; limited power budget means that only short-range links are often possible.

Traditional network communication frameworks such as link-state and distance-vector routing are less well suited to the problem of inter-satellite telecommunications because of the relatively low data-rate and limited transmission window which would be available due to the network topology [1]. The network topology is dependent on the constellation design, which itself depends on the application at hand. For example, Walker-Delta constellations consist of satellites which are generally optimally spaced to provide constant or near-constant coverage of the globe up to a specified latitude. A contrasting design is the flower constellation, which can be used for both local and global coverage [2]. Many existing and proposed telecommunications constellations (e.g. Iridium [3], Globalstar [4] and OneWeb [5]) consist of satellites in circular low Earth orbit (LEO), allowing each satellite regular, albeit possibly infrequent, contact with ground stations and in some cases for ISLs to be maintained permanently giving reliable and stable communications between different satellites. Crucially, a constellation of satellites in circular LEO benefits from periodic passes above the selected ground stations, therefore allowing a pre-defined link strategy to be used for the entire mission duration. For constellations of satellites placed into one or many different elliptic orbits, particularly if those orbits have very high apogee, the possible contact time each satellite has with a ground station is much reduced (due to the high distance to the ground station relative to the limited communication range), compared with the contact time available if it were in a circular orbit at low altitude, meaning that for effective communications with the whole constellation as well as Earth, ISLs are much more heavily relied upon. Additionally, the continuously-changing topology and mutual distances of the spacecraft require a continuous, real-time update of the ISL routing, with added reliability if this is performed on-board and not on Earth. Indeed, research has shown that for an application of a certain constellation of largely autonomous satellites involving distributed payload data, ISLs bring benefits to the constellation-to-Earth downlink capability, as they allow this data to be spread across the constellation [6]. Although making such heavy use of ISLs to communicate with Earth is currently rare [7], this would be especially important if the constellation described above were formed of CubeSats, due to their limited communication bandwidth and range. At present, communication, and in particular real-time communication with Earth would be challenging if at all possible in a constellation of CubeSats deployed into high-eccentricity orbits. One application of such a constellation would be in creating a distributed sensing platform which requires real-time data download to Earth from a small number of CubeSats at any one time [8].

It is clear that in a constellation such as this, the optimal communications solution should be able to adapt to the time-varying nature of the network topology while also being self-governing to a certain extent, to cope with the fact that each satellite may only spend a very small percentage of its orbital period in direct line-of-sight with a ground station. Previous research [9] has also identified that some proposed subsystem scheduling solutions are not possible for CubeSat constellations due to the additional constraints on energy-efficiency and the amount of on-board computation. These factors must be accounted for in order for the solution to be feasible. Although the scheduling of satellite subsystem operations in a constellation is a separate problem to telecommunications routing, it still requires an optimal solution to the usage of various distributed platforms in a satellite constellation, where each individual platform can be out of direct communications range of most others. It is therefore reasonable to apply the same conclusions to the problem discussed above to that of telecommunications routing optimization.

Previous research [10] has identified that as the control of assets in space is becoming more decentralized, dynamic routing algorithms which can direct communications according to the real-time network topology will allow for optimal solutions considering constraints such as path redundancy and platform power. Other research [11] has explored the routing strategies in large LEO and medium Earth orbit (MEO) constellations where there are enough satellites to create an ISL mesh. This solution is satisfactory for large circular LEO/MEO constellations, but for constellations consisting of different orbit eccentricities/altitudes and requiring many different ISLs to make up one communications channel, it is not optimal. The application of traditional networking techniques to the problem of inter-satellite telecommunications has had success [12] [13] but is not practical in large constellations where centralized decision making is not possible. It has been suggested in previous research that future focus in the field of reliable satellite networking should lie in efficient and optimal routing algorithms [14].

It has been identified [15] that autonomy is preferred for CubeSat swarms due to the low operational cost of such missions, where this autonomy is in the form of implemented behavioral rules which produce overall desired behaviors in the whole constellation. It has also been noted [16] that a heuristic solution is much preferred in the problem of satellite constellation networking than a more traditional deterministic approach because, as the algorithm is distributed, there is no single point of failure and is self-organizing so the solution adapts to the long-term variation in topology and node congestion. In this framework, each satellite is only required to be aware of the positions of the satellites immediately surrounding it, not of the full constellation. A route to a destination node cannot be computed on-board before transmitting data packets. It is therefore

necessary to make use of separate data packets which have the task of route-finding. These data packets are launched from a particular satellite with the purpose of finding appropriate ISLs which give a near-optimal route to Earth, and move between satellites independent of any other data traffic. Once a route is found, data would be able to flow along this route for a limited time, after which the topology of the network will have changed, breaking one or more of the ISLs. The routing data packets are mostly autonomous, so the nodes in the network only have to passively relay the routing packets after a simple calculation, using very little of their available power/processing time.

Previous studies which have examined ISLs as a means of increasing bandwidth in satellite network communications have placed strong emphasis on simulation as a means of truly understanding the effectiveness of theory [1] [10] [13]. Despite this, there has thus far been no case study or simulation of a complex satellite constellation to achieve meaningful results of the overall bandwidth of a large constellation. Previous research in this field has either focused on circular constellations in LEO/MEO [11] [13] or has specifically looked at high Earth orbit (HEO) constellation design [17] but considered only minimal ISLs in any route and so complex route-finding algorithms where communication links between two points are formed of many ISLs are not discussed.

The application of swarm intelligence to the problem of satellite constellation networking has been considered before [1] [16], with a conceptual focus and from the perspective of connectivity as opposed to obtaining tangible results with regards to data traffic flows and bandwidth. This type of solution was shown to be of comparable quality to existing ad-hoc routing algorithms, with the added advantage of being able to adapt well to significant changes in environment, whether that is the network topology or satellite failures. In the case of constellation networking, the density of available nodes in the network will vary significantly with time and location.

A swarm intelligence solution which has good performance in networks with highly time-variant topologies is found in ant routing algorithms [18] [19]. These algorithms mimic the collective behavior of ants looking for food. The purpose of these is to mimic the behavior seen in ant colonies when they are searching, as a group, for food sources. Each ant on its own has little intelligence, but many ants working together can achieve their common objective efficiently. The parallel with constellation networking is that route-finding data packets can be launched from any satellite, acting as 'scouts', and working together to ultimately find a near-optimal path through the other satellites (which are effectively relay nodes) to reach the destination node. The destination node would be one which can form a direct communication link with Earth. Ant routing algorithms allow good

performance in route finding because completely new routes can be explored without a heavy bias toward previous decisions. Research on the application of ant routing algorithms to inter-satellite telecommunications [1] has determined that one of the main reasons why these algorithms give good results is because the scheme is intelligent enough to enable self-organized re-routing according to communications requirements, based on decentralized coordination among the route-finding data packets. However the research conducted did not consider constellations consisting of multiple orbital planes. Similarly, although Refs. [20] and [21] also address the problem of telecommunications in satellite constellations by using the ant routing algorithm, the solutions are applicable to a LEO constellation of conventional-sized satellites. The problem of a high-eccentricity/altitude constellation formed of CubeSats presents different constraints and therefore requires an alternative solution. In addition to the differences in the spatial arrangement of the constellation, the data transmission method used in [20] does not satisfy a requirement for real-time data download. Additionally, traffic congestion is identified as being one of the main issues to be addressed with the ant routing algorithm, where data packets are buffered at popular nodes to be transmitted as soon as possible. Any buffering of this sort reduces the capability of the solution to provide real-time communications with certain points in the constellation. A solution which tackles the problem of congestion during the route-finding process and provides dedicated communications links which are available for limited periods of time is therefore required.

Ant optimization algorithms were also applied to satellite constellations in a different context: that of scheduling and allocating resources to the time windows available [22] and [23]. For example, in an Earth imaging constellation of satellites, some images are of higher monetary value than others depending on the mission, e.g. images of certain land masses or at certain latitudes are worth more than others. In these cases, the actuation of the imagers can be optimized such that the highest-value images are captured over the particular ground tracks of each satellite. Despite the use of ant optimization algorithms, the nature of the problem is different from the one under investigation; however, it shows the potential of ant optimization.

Thus far, the research undertaken specifically on ant routing algorithms has employed only straightforward constellations and scenarios on which to obtain results and has not focused on obtaining data in real-time from only a small number of nodes in the constellation. Furthermore, more detailed information such as nodal congestion and overall bandwidth of a distributed satellite system which would further determine the feasibility of these algorithms have not been explored in detail. The objective of this study therefore is to develop a strategy to maximize the available bandwidth of data download to Earth of a large constellation of satellites through optimal signal routing. By designing the algorithm to provide dedicated telecommunications routes

which are available for periods of time, the issue of packet buffering will be removed, allowing the data download to occur in real-time. A large constellation of CubeSats in high eccentricity/altitude orbits is considered, as this presents the greatest challenge because of the small percentage of orbital period which each CubeSat can communicate with Earth. The test case constellation is the CENTINEL [8] proposal which is a CubeSat constellation consisting of 74 CubeSats distributed among several high eccentricity/altitude orbits. The purpose of this mission is to collect space weather data at the Earth's magnetotail, surrounding the apogees of the highest-altitude orbits. With no telecommunications routing strategy in place, this data would have to be downloaded to Earth during the brief time where the satellites reach their perigee. The proposed approach is general and applicable to any constellation with little to no change. Ant routing algorithms will be used to maximize the available bandwidth in order to download as much data as possible in real-time from any point in the constellation to Earth. Although previous research has considered ant routing algorithms as a solution to satellite constellation telecommunications, this has not been done with the aim of achieving real-time data download from any point in a high eccentricity/altitude constellation.

II. METHODOLOGY

A. Ant System Algorithm

In order for direct communications links to be found between any two specific, pre-determined nodes in the network, a route-finding algorithm shall be designed. It shall be able to operate discretely, in any part of the network, independent of any central decision-making node connected to all other nodes in the network, as can be found in many other network routing problems. The proposed solution makes use of an algorithm known as Ant System© (AS) [19], and aims to mimic the behavior seen in ant colonies when they are searching for a food source. The way in which ants achieve this is an example of 'swarm intelligence' where a group of relatively unintelligent units work together to achieve a common objective. A brief description of how ant colonies use swarm intelligence to achieve their objective of finding food sources is summarized below.

Individual ants leave their colony in various random directions, depositing a constant trail of pheromone on their path. If an ant successfully discovers a food source, it follows its own pheromone trail back to the colony while still depositing more pheromone, see Fig. 1 [18]. This means that the path leading to the food source now has double the amount of pheromone on it. When another group of ants leave to search for food, they are more attracted to paths which have pheromone deposits from other ants. If two food sources are found and one is far away, then by the time the ant which found further away food source reaches the colony again, the path leading

to the closer food source will have been reinforced with pheromone many times over by several ants. Since the amount of pheromone on a path is proportional to the statistical likelihood that an ant will follow that path, this means that subsequent ants are much more likely to visit the close food source, but some small number will still visit the more distant source. At every stage there is always a chance that an ant may break off a pheromone path and find an even closer or bigger food source.
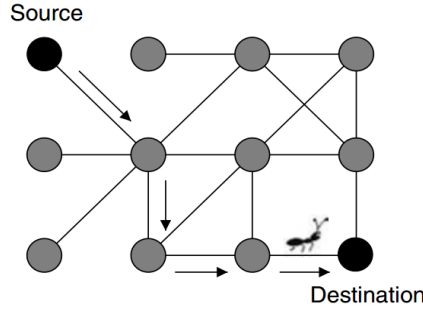

Fig. 1: Ants have various route choices between source and destination; their decisions when travelling from node to node is probabilistic.

This principle of many units scouting out routes and returning to a colony, having rewarded a path based on not only whether it was successful but also on its distance, and having route choices determined by a probabilistic calculation at each step, is what the AS algorithm is based on. AS was initially developed and tested on the travelling salesman problem (TSP). This is the problem where a salesman wishes to visit all cities on a map and needs to determine the shortest possible route to take.

The mathematical structure of the algorithm is well documented elsewhere [19], so only a summary is presented here. The most important concept in the algorithm is the fact that the route-finding 'ants' travel to their destination, making decisions on which nodes to travel to, based on a probabilistic calculation dependent on both the amount of pheromone leading to other nodes and a heuristic metric based on the amount of data traffic already flowing through each available node. The likelihood that an ant will go to any specific node depends on the node's probability compared to the others. The probability that ant k at node i visits node j is given by:

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha [\eta_{ij}]^\beta}{\Sigma_{l \in N_i^k} [\tau_{il}]^\alpha [\eta_{il}]^\beta} \ , \quad if \ j \in N_i^k \tag{1}$$

where τ_{ij} represents the pheromone of edge ij and is the proportion by which an ant is enticed to visit node j while at node i. η_{ij} is the heuristic factor which represents the desirability of each node based on the immediate environment only. The subscript l represents every other node other than nodes i and j. So if l is out of range of

i, then τ_{il} is zero. N_i^k is the set of nodes which an ant at i could possibly visit. α and β are the weights given to pheromone quantity and heuristic factor respectively.

The initial value of pheromone τ_0 (same for every edge) is calculated according to [18]:

$$\forall(i,j), \qquad \tau_{ij} = \tau_0 = \frac{m}{C^{nn}} \qquad (2)$$

where m is the number of ants per generation and C^{nn} is the length of a typical route. Initial pheromone must be set in this way so that the additional pheromone which is deposited at the end of every generation is in roughly the same order of magnitude.

The pheromone values τ_{ij} are updated after every generation of ants completes its exploration. Research has shown that updating pheromone in real-time as every ant progresses leads to poor results [19]. The pheromone update stage comprises of two steps:

1. Pheromone evaporation takes places across every non-zero pheromone value, whether it was explored by the ant or not, according to:

$$\tau_{ij} \leftarrow (1 - \rho)\tau_{ij} \qquad (3)$$

This allows any pheromone corresponding to poor routes to gradually fade away over successive generations.

2. Next, the pheromone level for the points making up any successful routes is increased, inversely proportional to the length of the route, according to:

$$\tau_{ij} \leftarrow \tau_{ij} + \sum_{k=1}^{m} \Delta\tau_{ij}^k \qquad (4)$$

where $\Delta\tau_{ij}^k = \frac{1}{C^k}$ when an ant actually travelled from i to j, C^k is the length of the full route. $\Delta\tau_{ij}^k = 0$ if an ant did not travel from i to j.

The algorithm iterates until a specified number of generations have been completed. If the algorithm converges, an optimal or near-optimal route should have been found. The reason for applying the AS algorithm for this purpose is for route-finding that is not centrally controlled: even if there is no guarantee that the actual optimal route is found at all times, any route which is close to optimal is sufficient.

B. Constellation Model

The model simulating the satellite constellation orbital dynamics was constructed. Keplerian orbital dynamics are employed to determine the true anomaly, and therefore position, of each satellite on its own orbit at any given time-step [24]. Although the relative positions of the satellites change with time, the entire constellation is assumed to be fixed at a particular time-step while the algorithm is being run. Routes should be explored, an optimal route converged upon and scientific data downloaded via this route in a reasonably short time in which the satellites do not significantly change position and so can be assumed to be fixed. Once enough time has passed for this assumption to be invalidated, the algorithm is re-started.

Each satellite can switch between three modes:

1. Transmitter Satellite (T_{sat}): a satellite generating and broadcasting raw data (for example, from a payload) which is to be downloaded to Earth. When a satellite is in this mode it is a source node for data packets;

2. Data-relay node: it receives data from one satellite and transmits it directly to another satellite in range;

3. Receiver Satellite (R_{sat}): a satellite in direct sight and range of a ground station, and therefore able to broadcast received data to Earth. When a satellite is in this mode it is a data-sink node in the network and a final target for data packets.

The number of available R_{sats} at any point in time varies considerably. The positions of ground stations are defined according to their latitude/longitude coordinates. The ground stations are fixed in an Earth rotating frame, and therefore rotate about the Earth's Z-axis at a rate of 360º per 24 hours in an inertial frame. Again, we assume that during each run, the rotation of the Earth is negligible, and hence the ground stations can be considered fixed in space and in the network topology.

1. Satellite Visibility

All satellites are assumed to use omni-directional antennae, e.g. the beam power density is the same in every direction. The maximum communication range of each satellite is assumed, in this work's test case scenarios, to be seven times the Earth's radius. In reality, bitrate is usually inversely related to distance, however we consider

a fixed bitrate when two satellites are within range, as this study is a proof of concept only. A constraint is added so that one satellite is visible to another (and thus a communication link is possible) when they are both within communication range and when the straight line connecting any two satellites does not intersect with the Earth.

2. Transmitter Satellites

In this work, the satellites in the constellation which actually generate the data to be downloaded to Earth in real-time are selected according to their true anomaly. This means that data will only be collected and transmitted in specific arcs of the orbit (those crossing the magnetotail, near the apogee), allowing for continuous or near-continuous sensing of these areas. Fig. 2 shows that the apogee of the orbit aligns with the magnetotail of the Earth. Individual satellites behave as data-relay nodes for the majority of the orbit, until they reach a specific arc close to the apogee. For the time that a satellite is on this arc it changes mode to become a T_{sat}. In Fig. 2, the red arc (bold) is identified as an area of interest in the Earth's magnetotail, satellites are represented by dots/cross and the color density plot represents Earth's magnetic field. It is worth underlining, however, that any other mission-specific criteria could be chosen to identify the T_{sat}.

3. Receiver Satellites

The Receiver Satellites are those able to communicate directly with a ground station, and hence at the end of the data-link. Any satellite which is both within communications range of the selected ground station(s) and is above the local horizon of the ground station(s) can be identified as an R_{sat}. The maximum possible bandwidth of the full system is the data transmission rate of one satellite multiplied by the number of R_{sat} so the number of these satellites at any given time is very important.

With a simulation model containing these constraints in place defining the start and end points of the data link as well as the visibility of each satellite in the constellation to every other, it is possible to apply the route-finding algorithm.

We have included supplementary MP4 video files which show features of this constellation model in motion. Please refer to the files *yarr_ceriotti_v1.mp4*, *yarr_ceriotti_v2.mp4*, *yarr_ceriotti_v3.mp4* and the associated descriptions, available at http://ieeexplore.ieee.org.
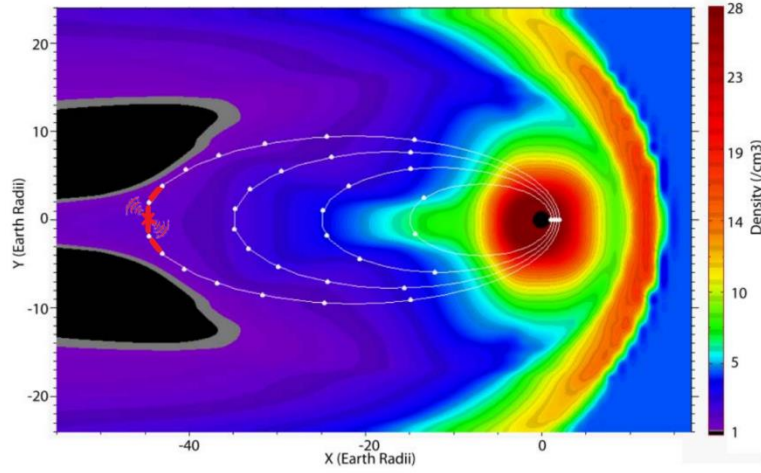
Fig. 2: [8] Constellation of satellites, with the arc around the apogee of the outer orbit identifying the transmitting region. When a satellite enters this arc it changes its mode from data-relay node to T_{sat}, collecting and broadcasting scientific data.

C. Ant System as a Solution to Constellation Networking

The AS algorithm can be applied to the simulation environment described. In the satellite networking problem, the ants would take the form of small data packets which can be passed to different nodes and which keep a log of each node already visited. The TSP itself is mostly comparable with the problem of route-finding in a satellite network save for two distinct differences. In the TSP the route is permitted to begin at any point, in satellite networking routes can only begin at pre-determined T_{sats} and can only end at pre-determined R_{sats}. This is not a trivial difference because in the travelling salesman problem ants can start and end at various points and all of their journeys will contribute toward the final solution, so many more parts of the network will be discovered per group of ants compared to satellite networking. Furthermore, ants departing from different T_{sats} develop pheromone trails independently of ants from other T_{sats}. The other difference between the two problems is that in the TSP, the solution must be a route which visits every node in the network, which is not the case in the problem of the satellite network problem. Although both problems are route optimization problems, constraining the start/end points of the route and removing the requirement that all nodes be visited changes the algorithm required to achieve a near-optimal route. Fig. 3 shows an illustration of these differences.
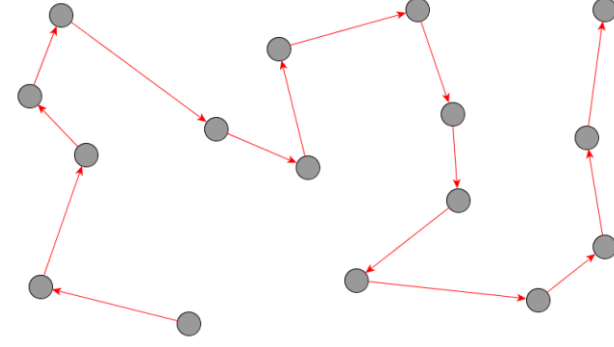
(a)
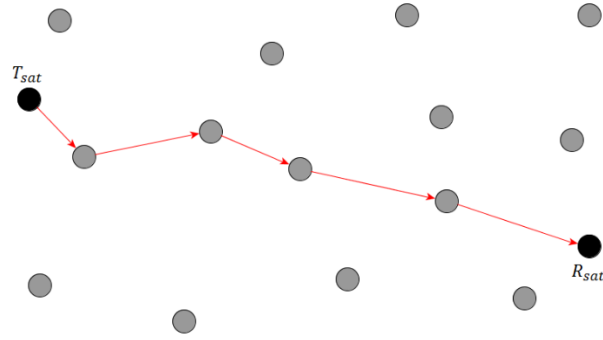
(b)

Fig. 3: (a) shows an illustration of a solution of the TSP: all nodes are visited and the start/end nodes can be anywhere. (b) is an illustration of the satellite telecommunications problem: the most optimal route must be found between two pre-defined nodes.

As in the application of this algorithm to the TSP, τ_{ij}, in Eq. (1) is the pheromone enticing an ant at node i to hop to node j. The heuristic factor η is taken as the inverse of the available bandwidth at each satellite, however in this application it is a vector, with η_j the inverse of the bandwidth of node j, which is any node which the ant could visit next. This means that satellites which have less bandwidth available will be less likely to be visited by ants than those with more. By including this extra factor in the probability calculation which the ants perform at each node, it ensures that routes are selected based on practicality as well as length. The probability calculation in Eq. (1) must therefore be modified to suit this application, and is:

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha [\eta_j]^\beta}{\Sigma_{l \in N_i^k} [\tau_{il}]^\alpha [\eta_{il}]^\beta} \, , \quad if \ j \in N_i^k \tag{5}$$

For the pairs of nodes within range of another, the initial pheromone is set to τ_0. For every other node, it is set to 0.

When good data links are established between T_{sats} and R_{sats}, some ISLs will eventually be lost as the network topology changes. However, a certain small amount of bandwidth can be reserved in each satellite for the exclusive use of route-finding packets, so the route-finding process can be in operation continuously meaning there are no breaks in data flow. Route-finding packets are essentially the ants in the AS similitude, and will be referred to as ants in the following. The ants are given priority over actual data packets, so new communications routes could be being sought while old ones are still in use. This allows continual adjustment of communications to cope with the changing positions of the satellites and the revolution of the Earth. This also makes communications self-healing as if a node fails completely for whatever reason, communications will not be disrupted to a large degree because routes are continually being updated. The failed node would then correspond to a row of permanent zeros in the pheromone matrix.

When an ant reaches an R_{sat}, it then re-traces its route back to the original T_{sat}. For clarity, and in-keeping with the original algorithm [18], the term ant comprises of a forward and a backward route-finding packet. At each generation, a T_{sat} releases a number of forward-ants simultaneously, each of which moves independently from node to node according to Eq. (5) while a log of each node visited is stored within it (each CubeSat would have a unique identification). If the forward-ant reaches the R_{sat}, the log is passed to a backward-ant and the forward-ant is then immediately deleted. The backward-ants then visit each node in their log in reverse order, updating the pheromone values at each link, according to the quality of the total route, as in Eqs. (3) and (4). When all backward-ants have returned to the original T_{sat}, the same process begins again with a new generation of forward-ants; this time the route selection of the new generation will be influenced by the results of the previous. As route selections are based on a probabilistic calculation, this algorithm is very much a stochastic process.

The entire $n \times n$ pheromone matrix is not itself stored anywhere in the network as a whole, but is fragmented across the entire network, so in effect each satellite possesses one row of that matrix. After a pre-specified number of generations, actual scientific data leaves the T_{sat} and chooses the node with the highest pheromone value at each step, which should bring it to an R_{sat} as quickly as possible. Each satellite would also need to have a pre-set time interval after which the route-finding process would begin again from scratch, so at that point all satellites reset their pheromone values to initial conditions, and the process begins again. The number of generations before actual data is sent must therefore be pre-determined, along with the time interval between pheromone matrix resets.

1) Optimization Parameters

In addition to tuning the AS algorithm, different conditions for terminating a forward-ant can be studied. The route log created by a forward-ant can be used to define different route permissions which the forward-ants must follow. Three different route permissions are examined in this study, and illustrated in Fig. 4:
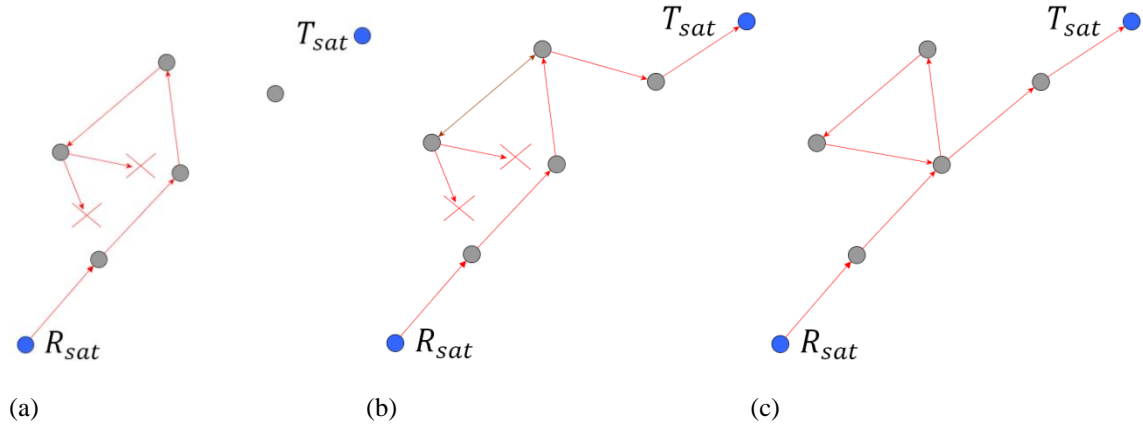

Fig. 4: a), b) and c) illustrate route permissions 1, 2 and 3 respectively.

1. Each satellite can only appear in the route log once. This means that if a forward-ant has at any time the only option is to revisit one node, ant is terminated immediately and no backward-ant is launched.

2. If the ant finds itself in the same condition as above, a backward step is taken to the last node on its route log, and delete the dead-end node. This is done simply by setting the relevant pheromone value to zero. The ant then continues.

3. The ants can visit any node, regardless of whether it already appears on its route log. With this method, every ant is always guaranteed to reach the destination node however the route log will likely contain a large amount of loops.

As well as changing the route permissions, the other condition to put in place for a forward-ant to terminate, is to limit its lifetime, i.e. the number of nodes it can visit before it terminates, without launching a backward-ant.

Once this algorithm has been designed and has been shown to produce the shortest possible, or near the shortest possible route between two satellites in a constellation, the wider question remains of how to maximize the bandwidth used when multiple data links are to be found between outlying satellites and Earth. There are therefore two scenarios to be considered: 1) finding a single route between two pre-defined nodes, and 2)

optimizing this method for the case of multiple start and end nodes, to maximize the total bandwidth achievable with a particular topology of the constellation network.

We have included a supplementary MP4 video file which shows a visualization of this algorithm in progress. Please refer to the file *yarr_ceriotti_v4.mp4* and the associated description, available at http://ieeexplore.ieee.org.

2) Reducing Congestion

In the constellation, there will be more than one T_{sat} which need to download data to Earth, and may be in different positions within the constellation. If forward-ants from each T_{sat} followed Eq. (5) while sharing the same pheromone matrix, the final route choices would be sub-optimal as the value of a particular edge *ij* will be different for each T_{sat}. There should therefore be multiple pheromone matrices, each corresponding to a particular T_{sat} in order to achieve the most optimal results. Each T_{sat} implements one route-finding algorithm at the same time as, but independently of, the others.

If multiple route-finding algorithms are being run simultaneously from different T_{sat}, there is a significant risk that many data packets will be assigned to one node when they could easily be spread out over many. This inefficient use of nodes eventually results in some of them being unable to relay data packets as soon as they are received, which creates gaps in the system making it more difficult for future routes to be found. This will be referred to as congestion, which has the ultimate effect of reducing the bandwidth of the full system. It must be stressed that depending on the T_{sat}/R_{sat} ratio and the constellation under consideration, some amount of congestion is inevitable. Where this is the case, it is still desirable to ensure congestion is kept to a minimum and for it to occur only in the certain places where it is unavoidable.

Fig. 5 shows a situation where multiple T_{sats} have found routes using one node in particular, while other nodes are available to some T_{sats}. If these other nodes were utilized, it would spread out the routes thereby increasing the bandwidth overall.
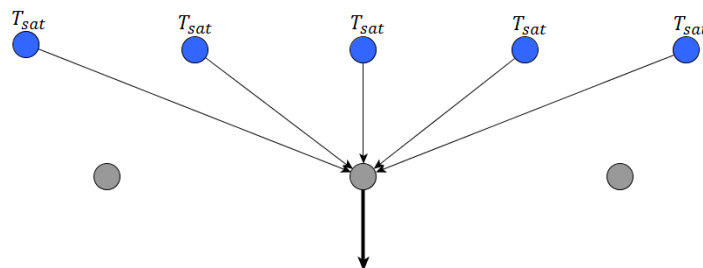

Fig. 5: Unnecessary congestion. Multiple T_{sats} using routes passing through a particular node, where other nodes are available.

To reduce congestion, one needs to determine which nodes are being used by which T_{sat} and, crucially, to what extent. If a particular node has high pheromone values corresponding to one T_{sat} but low for others, then additional priority must be given to the T_{sat} using the node most. This is done by evaporating extra pheromone which may lead forward-ants from other T_{sats} to this one, but only for those T_{sat} which have been found to be scarcely using the node. This additional pheromone evaporation will take place at set intervals as the generations progress, as the prioritization takes place on trends observed at the end of a set of generations. The prioritization takes place according to:

$$[\tau_{ij}]_x \leftarrow \begin{cases} 0.1[\tau_{ij}]_x & : [\tau_{ij}]_x < 0.8[\tau_{ij}]_m \\ [\tau_{ij}]_x & : [\tau_{ij}]_x \geq 0.8[\tau_{ij}]_m \end{cases} \quad (6)$$

where $[\tau_{ij}]_x$ is the pheromone enticing an ant launched from T_{sat} x at node i to visit node j and T_{sat} m is the T_{sat} which has the highest pheromone value at τ_{ij} at that time.

III. CASE STUDY

In order to simulate and assess the performance of the networking algorithm, a test case has been selected. The test case is based on the satellite constellation outlined in the mission proposal CENTINEL [8]. The purpose of this mission is to predict space weather events by locally sensing the magnetic field in the Earth's magnetotail continuously, by deploying a large number of satellites onto four orbit-pairs, where the semi-major axis varies with each pair, creating an arrangement where some orbit pairs are 'nested' within others. The only difference between the two orbits on each pair is the inclination, which is +/-18.5° for each. The semi-major axis and eccentricity of each orbit-pair is given in Table I, along with the mean anomalies of each satellite. Three active ground stations are arbitrarily selected for the analysis, their coordinates are given in Table II; at $t = t_0$ the ground station Cartesian coordinates are also given, where the origin is the center of the Earth, the positive X-axis extends in the direction of the orbit perigees and the positive Z-axis extends in the direction of the North pole. All distances are normalized to the radius of the Earth. A visualization of the constellation is given in Fig. 6. By deploying multiple satellites on each of these eight orbits, a distributed sensing platform is created, where those on the apogee of the orbits, which are in the magnetotail of the Earth (and would be far beyond communication range of Earth) can collect data.

Table I
ORBITAL PARAMETERS OF THE CENTINEL SATELLITES

Semi-Major Axis [km]	Eccentricity	Number of Satellites Per Orbit	Mean Anomaly of Each Satellite [°]
51342	0.8634	3	0 / 120 / 240
84189	0.8939	7	0 / 51.43 / 102.86 / 154.29 / 205.71 / 257.14 / 308.57
117036	0.9074	11	0 / 32.73 / 65.45 / 98.18 / 130.91 / 163.64 / 196.36 / 229.09 / 261.82 / 294.55 / 327.27
149883	0.9149	16	0 / 22.5 / 45.0 / 67.5 / 90.0 / 112.5 / 135 / 157.5 / 180 / 202.5 / 225 / 247.5 / 270 / 292.5 / 315 / 337.5

Table II
COORDINATES OF GROUND STATIONS AT $t = t_0$

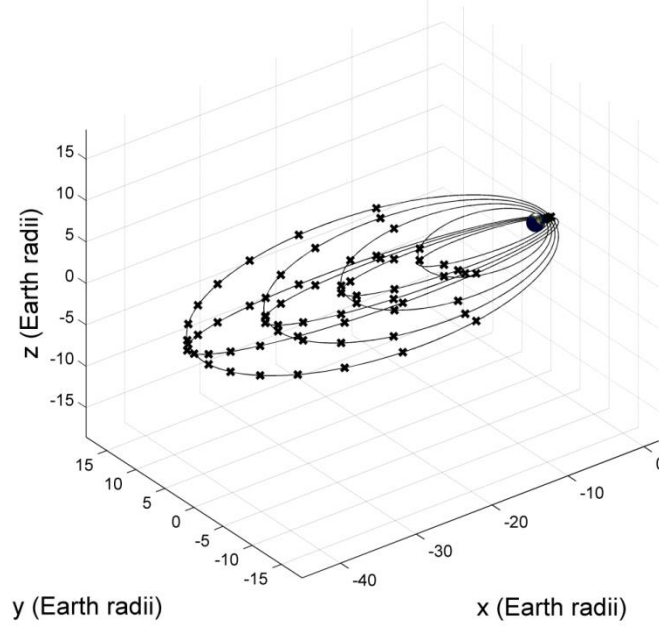| Ground Station	X-coordinate	Y-coordinate	Z-coordinate	Latitude [deg]	Longitude [deg]
Canberra, Australia	0.7054	-0.4242	-0.5678	+ 35.43	- 116.89
California, USA	0.3458	0.7378	0.5797	- 35.40	+ 148.98
Glasgow, UK	-0.5599	0.0363	0.8278	+ 55.86	- 4.26

Fig. 6: The constellation of 74 satellites on eight nested orbits which comprise the CubeSat constellation proposal CENTINEL [8].

Due to the large number of satellites required for a mission of this scale, the mission designers have specifically identified that CubeSats or a similar small-satellite platform must be used for this mission for it to be financially feasible. However, without ISL, data cannot be downloaded to Earth until each satellite is at or near the perigee of their orbits, making this a store-and-forward constellation. With routed ISL, satellites can collect and broadcast data in the region near the apogee, and the core within the constellation can be used as data-relays to create a pseudo-real-time link to the Earth. Thus, space weather information from the Earth's magnetotail can be collected in quasi-real time, providing data to detect and even predict, for example, space geomagnetic storms.

This constellation design suits this study well as it makes use of high-eccentricity/altitude orbits which take the CubeSats far out of range of Earth when they are in the area of interest, and the same orbits have low perigee, allowing satellites to be used as data-relays to ground stations.

IV. RESULTS

In applying the AS algorithm to the problem of constellation networking, the effects of several variables and constraints need to be examined to ensure that it will reliably provide a near-optimal solution with efficient use of computational power on each execution. There are two scenarios to be analyzed: (A) implementing and optimizing the algorithm to find a route between a single T_{sat} and a single R_{sat} and (B) further optimizing the method for the real-life scenario of multiple T_{sat} and multiple R_{sat} to maximise the bandwidth of the constellation at whatever the network topology it has at a given time-step.

A. Single-Route Results

In this test case, we consider a single T_{sat} and a single R_{sat}, as shown in Fig. 7, which have the orbital parameters shown in Table III. The constellation is at the time $t = t_0 + 35\ hours$, the coordinates of the ground stations at this time are given in Table IV.
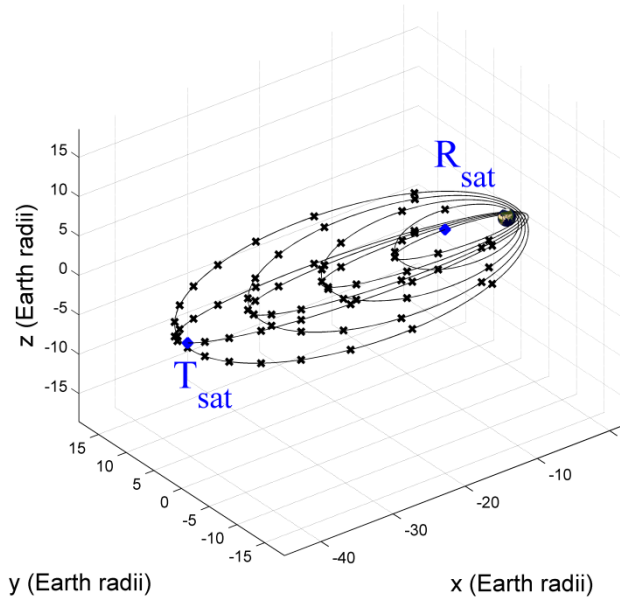

Fig. 7: Spatial configuration of constellation for single-route test case analysis.

Table III
ORBITAL PARAMETERS OF SINGLE T_{sat} AND R_{sat}

	Semi-Major Axis [km]	Eccentricity	Inclination [°]	Mean Anomaly [°]
T_{sat}	149883	0.9149	18.5	191.05
R_{sat}	51342	0.8634	-18.5	31.78

Table IV
COORDINATES OF GROUND STATIONS AT $t = t_0 + 35\ hours$

Ground Station	X-coordinate	Y-coordinate	Z-coordinate
Canberra, Australia	-0.5716	0.5923	-0.5678
California, USA	-0.5250	-0.6232	-0.5797
Glasgow, UK	0.5314	-0.1800	0.8278

To measure the efficiency of the method when applied to the problem of inter-satellite telecommunications, several performance indicators are assessed. These are:

- The length of the best route both in terms of number of nodes visited and actual distance in space from T_{sat} to R_{sat};

- The number of hops which occur throughout the entire length of the algorithm, where one hop represents one ant moving from one satellite to another.

- The number of ants which reach the destination R_{sat} in every generation.

- The number of generations it takes before ants converge on one route or a small group of near-identical routes.

One of the main difficulties in implementing AS for the problem of constellation networking is in striking a balance between allowing bias to develop to good routes (convergence) while still giving ants sufficient possibility to explore new routes (exploration). If the pheromone updates for promising routes are comparatively too high, future ants will keep exploring these routes, in turn reinforcing the pheromone even more, while ignoring alternative (and possibly better) routes. Conversely, if pheromone updates for promising routes are comparatively too small, future ants do not see any distinction between the good and bad routes explored previously. This means a balance is to be found between the three variables τ_0, ρ, and $\Delta\tau_{ij}^k$. Ideally, the algorithm should return a good route which is found after many generations of ants.

For the test case specifically, the value of τ_0 is set according to Eq. (2) where the value of C^{nn} is what would be considered a good route length, for a route starting near the apogee (2.5×10^5 km), giving $\tau_0 = 0.25$. The value of $\Delta\tau_{ij}^k$, which is updated by the backward-ants as they return to the original T_{sat}, is set according to Eq. (4), and is the reciprocal of each successful route length. Although an optimal value of $\rho = 0.5$ has been

suggested [18], this is for the application of this algorithm to the TSP. There are significant differences between this and the problem at hand as explained previously, so a good value for ρ must initially be sought. Setting $\rho = 0.5$ as advised in [19] does not produce optimal results. In the TSP, ants pass through every node with each trial, so pheromone updates are being applied to nodes very regularly, which justifies the higher pheromone evaporation rate. Since fewer parts of the network are being discovered per generation of ants compared to the travelling salesman problem, evaporating pheromone at this rate is too aggressive. In the test case, it is likely each ant does not visit the majority of the nodes available, so clearly the pheromone evaporation rate should not be set to as high a value as in the TSP. After investigation it is found that the optimally tuned value for the test case is $\rho = 0.08$. This value is chosen because it consistently results in a short route length each time the algorithm is run. Although other values can regularly result in routes of similar length, this value has the smallest standard deviation of route length across a number of runs and is chosen for that reason. This value is used in all subsequent simulations, and should be tuned according to the constellation under consideration.

The values of α and β, which are the weighting between the pheromone and heuristic factors in Eq. (5) are both set to 1 at this stage. The relative values of these variables is more important when considering multiple routes in the same constellation.

Several trials are conducted, with the purpose of finding a route between one particular T_{sat} close to the apogee of the largest orbit, and one R_{sat}. 200 generations are used per trial, with 10 ants per generation, so a total of 2000 ants are launched in each trial. The routing permissions at this stage are set as in Section II.C.1), and there is no limit on ant lifetime. The objective is to obtain a route which is as short as possible, and the only termination condition is on number of generations.

Fig. 8 and Fig. 9 show the progress of the algorithm at each generation, for route permissions 1 and 3 respectively, with the AS parameters optimally tuned. Fig. 8(a) shows percentage of ants in each generation which reach an R_{sat} before being terminated, and hence launch backward-ants to update pheromone. Panels (b) and (c) show a similar trend to convergence for both the length of the routes found and the number of hops used during the full algorithm. Note the number of hops disregards whether the forward-ant successfully reaches the R_{sat} or not. Convergence is defined as the point where all subsequent forward-ants always follow one path to the R_{sat} or a path very similar (i.e. a path where only a very small number of ISLs are different). In Fig. 9, the percentage of ants in each generation reaching an R_{sat} is not shown as with this route permission they are all guaranteed to reach one. Due to the stochastic nature of the algorithm, each of the three route permissions are simulated 10 times each, and the metrics described above are averaged and presented in Table V. The number of

nodes used does not include the T_{sat} and R_{sat} and the length of the best route found is the sum of the distances of all the ISLs used in the route.

(a)   (b)   (c)
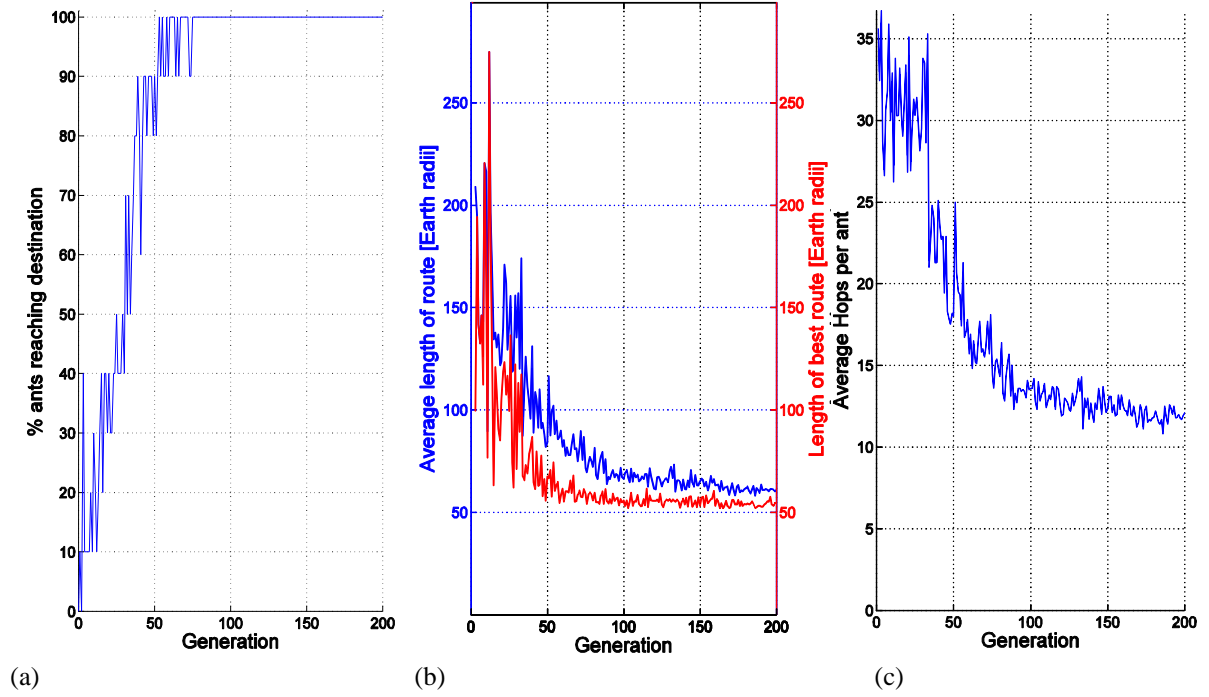
Fig. 8: Algorithm progression with route permission 1. (a) number of ants which reach the R_{sat} without being terminated. (b) blue line: average length of routes; red line: length of the shortest route. (c) average number of hops taken by each ant in each generation.

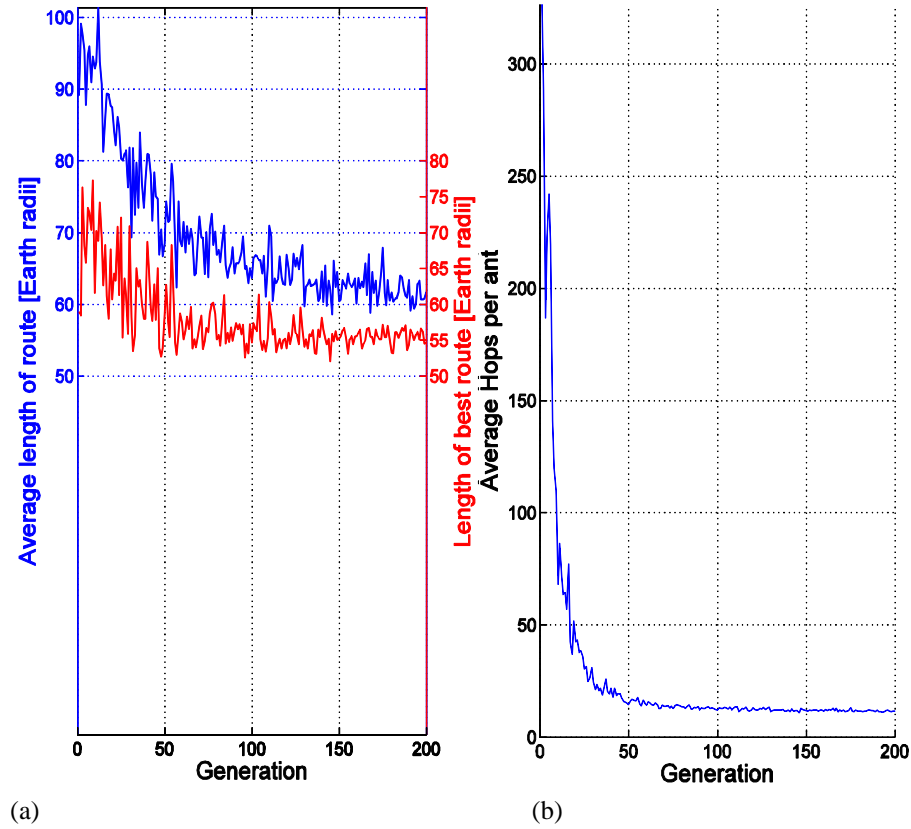(a)   (b)

Fig. 9: Algorithm progression with route permission 3. (a) blue line: average length of routes, red line: length of the shortest route. (b) average number of hops taken by each ant in each generation.

Table V
AVERAGE OF THE RESULTS OF 10 SIMULATIONS OF THE ROUTE-FINDING ALGORITHM RUN WITH DIFFERENT CONSTRAINTS.

Route Permission	Avg. nodes used in best route found	Avg. length of best route found [Earth radii]	Avg. hops during full run	% ants reaching R_{sat}	Avg. number of generations until convergence
1.	9.1	52.251	34616	86.47	154
2.	9.2	51.269	41367	100.0	125
3.	9.0	50.730	49687	100.0	67

Comparing the results generated for the three route permissions in Table V, it appears that route permissions 2 and 3 give a better average length of the best route found than permission 1. This is because in both methods, every single ant is guaranteed to reach the destination R_{sat} which makes the learning process more efficient. For these two permissions, at the end of every generation, all 10 routes found will result in pheromone being deposited because every forward-ant will have reached the R_{sat} and launched a backward-ant. This is not the case with permission 1. as some forward-ants are terminated before any backward-ant is launched causing less pheromone to be deposited for each generation, which in turn causes the general learning process to progress more slowly. Although enabling loops (as in route permission 3) results in the shortest average length of the best route found, the early generations of ants can create routes which are an order of magnitude longer than those in the other two settings, as seen in Fig. 9. This explains why this method results in the highest total number of hops. Before pheromone is deposited when loops are allowed, the loops which appear in the route are deleted. This means that with route permission 3 enabled, ants can generate good routes early in the exploration.

Based on these results, route permission 3 is deemed the ideal setting, as on average this results in the shortest routes with a relatively small time to convergence. However, when this algorithm is further adapted for use in a full constellation networking scenario with multiple T_{sats}, it must be adaptable to network congestion. The very small number of generations to convergence observed when route permission 3 is enabled does not allow for this. Therefore, the optimal permission for this application is permission 2. (backtracking allowed).

The effect of limiting the lifetime of forward-ants was also investigated: it was found that, although this will inevitably result in a small number of them not reaching their destination, the main effect is to reduce the average number of hops in the run. With constrained lifetime, the algorithm can repeatedly produce an optimal or near-optimal route between two satellites in a constellation using minimal total hops. An example is shown in Fig. 10, where each dot represents a satellite and the T_{sat} is near the apogee of the highest eccentricity orbit-pair. Data is transmitted from the T_{sat} and selects the next node to move to at each step by choosing the one which has the highest level of pheromone, the solid line therefore represents the ISLs which are formed at the end of

the algorithm. The route ends at an R_{sat}, so data can be downloaded directly to Earth. The route-finding algorithm was run 100 times and each time the ants converge either exactly on the route shown, or on a very similar one.
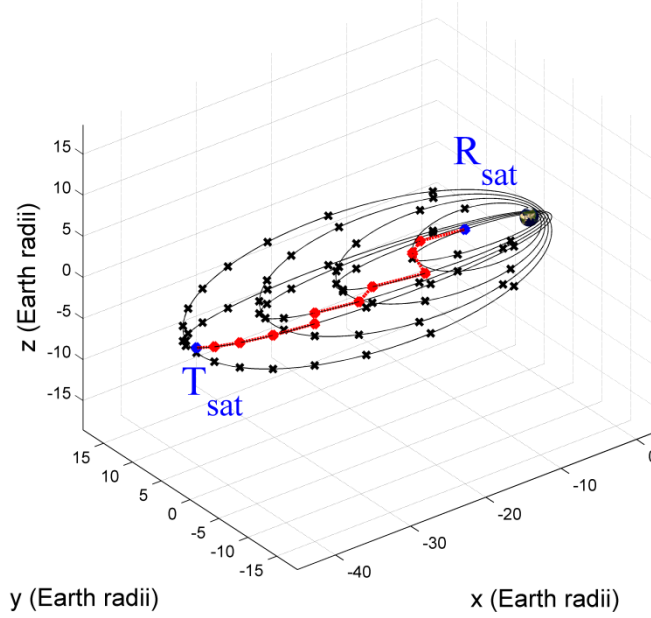

Fig. 10: An example optimal route. T_{sat} and R_{sat} are highlighted in blue.

B. Constellation Results

With the AS algorithm successfully implemented to get repeatable and short routes between two satellites in the constellation, the constellation networking problem can be addressed, where multiple satellites act as T_{sat} and R_{sat}. The route-finding algorithm must be integrated into a full solution for real-time downloading of data to Earth which maximizes bandwidth. Of particular interest is how the resulting bandwidth of a system using the route-finding method outlined compares to the maximum possible bandwidth determined by the number of R_{sats} as well as the amount of data actually being transmitted by the T_{sats}.

If the purpose of a satellite constellation is the real-time downloading of data to Earth, then there will be several T_{sats} generating and transmitting data and several R_{sats} broadcasting to Earth. The maximum achievable bandwidth of this system is therefore the data transmission capacity of one R_{sat} multiplied by the number of R_{sat} available at any time. This maximum theoretical bandwidth may not be possible in practice, if congestion occurs in crucial relay-nodes in the network.

If each T_{sat} establishes one route at every instant in time, the bandwidth of the full system will be affected by the amount of congestion. Congestion in this context is defined as the bandwidth of a node being reduced due to multiple T_{sats} using routes which pass through it as illustrated in Fig. 5. In order to avoid congestion, the common networking approach of splitting data into packets which are re-assembled at the destination can be

taken. The normal method of packet switching is not the most ideal for real-time communications because of the unknown amount of buffering which can take place [25]. Instead, each T_{sat} will split its data into 10 separate packets, and therefore execute the route-finding algorithm 10 times, creating a dedicated route for each packet. This means that in an instance where there exists 6 T_{sats}, there will be a total of 60 routes to Earth, 10 for each T_{sat}, as opposed to 6. This will improve the overall bandwidth of the system by utilizing more of the bandwidth existing across all the nodes without causing the computation time to become unpractical.

This does not prevent congestion completely: if certain nodes in the constellation are heavily used due to their location, then they may reach their bandwidth capacity. If data cannot be stored, this means that these packets will be lost, and the actual bandwidth (as opposed to the maximum achievable bandwidth) of the system is reduced.

1) Tuning of the Probability Parameters

One tuning parameter to consider in optimizing available bandwidth is the weight given to the two factors in the probability calculation at each node, see Eq. (5). α is the weight assigned to the pheromone component and β is the weight assigned to the heuristics. The effect of the weights of these two factors on the final route selection is now investigated. The algorithm will run until a pre-defined number of generations has been launched, at which point actual scientific data will be passed along the route with the highest pheromone at that time. The amount of time taken before the algorithm is completed is of interest as this is directly influenced by the value of β. For this investigation, α is held constant at 1 while β is varied. The constellation is at time $t = t_0 + 46$ hours, and at this time there are 4 R_{sats} available. Three different arcs are selected on the two highest-eccentricity orbits where T_{sat} should be active, giving a total of 6 T_{sats} for this test case. Table VI displays the orbital parameters of each T_{sat} and R_{sat}.

Table VI
ORBITAL PARAMETERS OF T_{sat} AND R_{sat} BETWEEN WHICH TELECOMMUNICATIONS ROUTES MUST BE ESTABLISHED TO ALLOW REAL-TIME DATA DOWNLOAD TO EARTH.

	Semi-Major Axis [km]	Eccentricity	Inclination [°]	Mean Anomaly [°]
T_{sat} 1	149883	0.9149	18.5	103.23
T_{sat} 2	149883	0.9149	18.5	193.23
T_{sat} 3	149883	0.9149	18.5	283.23
T_{sat} 4	149883	0.9149	-18.5	103.23
T_{sat} 5	149883	0.9149	-18.5	193.23
T_{sat} 6	149883	0.9149	-18.5	80.73
R_{sat} 1	51342	0.8634	18.5	34.92
R_{sat} 2	51342	0.8634	-18.5	34.92
R_{sat} 3	84189	0.8939	18.5	348.08
R_{sat} 4	84189	0.8939	-18.5	348.08

As there are 72 satellites in the entire constellation, this leaves only 62 left to be used to find 60 routes to Earth, many of which are unusable due to their particular locations. As described above, each route will be used for the transmission of one tenth of the broadcast bandwidth of each T_{sat}. It is clear that it is not possible to find 6 completely independent routes which do not share any nodes, so congestion is (deliberately) inevitable in this test case. It is assumed that each T_{sat} transmits data at a rate of 5 Kbps and every other satellite in the constellation can relay data at a rate of 10 Kbps. This means that the total amount of data being transmitted is 30 Kbps and the total amount of data which is able to flow to Earth via the 4 R_{sat} is 40 Kbps.

Fig. 11 shows that as β is increased, the achievable bandwidth increases to a maximum, then decreases slightly. The time to completion increases overall. Although the number of generations is always limited and is what defines the completion of the algorithm, the time to completion will vary because the amount of time which a single ant takes to find a route varies. If several generations are needed to reach convergence, then the processing time will be longer. This is because the higher the value of β, the less the pheromone deposits contribute to the probability calculation. This means that forward-ants are increasingly making choices based mainly on their immediate environment with little regard to the successes and failures of previous ants. The effect of increasing β is to initially improve bandwidth, because more importance is given to avoiding busy nodes. The ideal value should therefore be set high enough that bandwidth is maximized, but not so high that processing time becomes significantly worse. A good compromise is found at $\beta = 3$ with $\alpha = 1$. These values are in good agreement with what was found for the TSP, where the ideal weighting was found to be $\beta = 5$ with $\alpha = 1$ [19]. The maximum bandwidth achieved with these conditions is 29.55 Kbps.
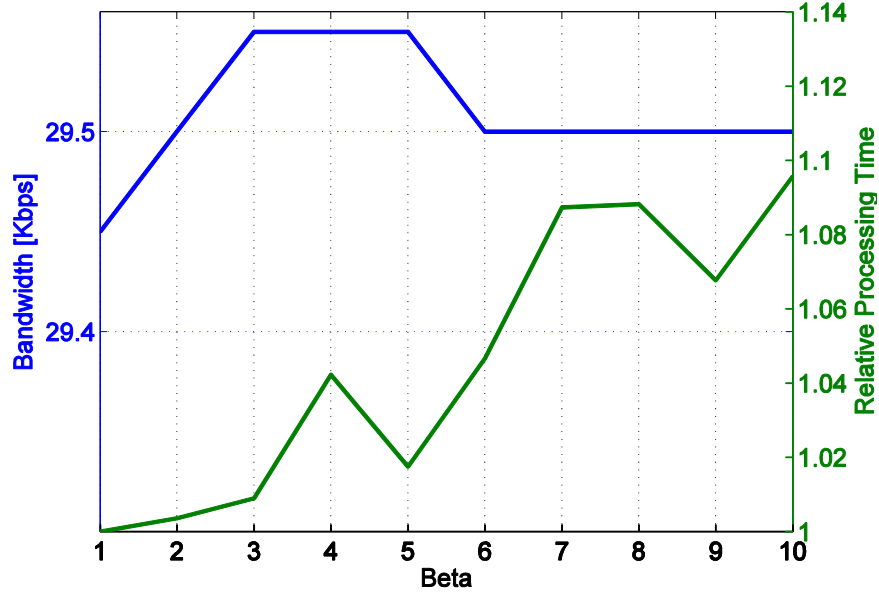
Fig. 11: Overall bandwidths achieved varying the weight.

2) Congestion

The problem here is to ensure that running the route-finding algorithm simultaneously from different T_{sats} will minimise packet congestion and therefore make the most efficient use of what bandwidth is available. This is illustrated in Fig. 5, where multiple T_{sats} are attempting to use one node when others are available causing it to become a bottleneck which then limits the real-time transmission rate achievable by the T_{sats} until the network topology changes. The problem with congestion as it has been defined here is that it must be dealt with intermittently during the route-finding process i.e. it cannot simply be added in as a heuristic factor in the probability calculations [19]. The heuristic factor described by the matrix η represents the data transmission capacity, which discourages ants from visiting nodes which have already been assigned many data packets from the previous route-finding process, but this does not account for the congestion which is caused by multiple T_{sats} attempting to create routes through the same node in the next route-finding process, where this is not necessary.

An anti-congestion mechanism for implementing the analysis/prioritization described in Section II.C.2) is included in the search algorithm. This is essentially a route prioritization process. Now, unless the pheromone matrix has been updated a number of times, the system will not have learned enough since the previous update. The prioritization process will start to work when pheromone trends have started to be established. For this reason, the process is initiated after a set of generations has been completed (in this test case, it is after every 8 generations). After each set of generations, the different pheromone matrices corresponding to each T_{sat} are updated according to Eq. (6). In the real implementation, the matrices are dispersed across the entire

constellation so this analysis would take place at each node, where each node has access to one row of each pheromone matrix. This would require time-synchronization across all satellites.

The algorithm is tested using the same test case described above. Two settings are examined: one where there is no attempt to prevent congestion and another where the anti-congestion mechanism is used. The algorithm is run 10 times for each setting and the resulting congestion on all nodes is recorded, averaged, and shown in Table VII. A congested node is one where more than one route has been assigned to pass through it, at the end of the route-finding process. The amount of congestion is defined as the total number of excess routes which pass through nodes with multiple routes assigned, i.e. the number of routes a node is part of, minus one. If a particular node is part of 3 routes, this adds 2 to the total amount of congestion in the constellation. The total amount of congestion is therefore the sum of all congestion values of all the nodes.

Table VII
RESULTS OF ENABLING THE ANTI-CONGESTION MECHANISM. RESULTS AVERAGED OVER 10 RUNS

Anti-congestion mechanism setting	Avg. congested nodes	Avg. amount of congestion (no. of additional routes passing through nodes)
Disabled	14.1	28.4
Enabled	9.8	21.7

A qualitative measurement of the success of this mechanism is the observation of the routes selected at the end of the process. Fig. 12 and Fig. 13 show examples of the routes with the anti-congestion mechanism disabled and enabled respectively. The solid turquoise dots represent the congested nodes, having more than one route pass through them, while the red nodes have only one route passing through them.

Table VII shows that using the anti-congestion mechanism reduces the congested nodes by 30.5% on average and the total amount of congestion experienced by all nodes by 23.6% on average. This is a considerable reduction, especially considering that some degree of congestion is unavoidable. The mechanism reduces congestion to some degree, but also moves congestion away from where it can be avoided. The highlighted area of Fig. 12 shows that congestion may unnecessarily occur. The routes starting from the two T_{sats} in that region do not need to share the same nodes, but with no mechanism to prevent this, congestion might happen. The same region in Fig. 13 shows more favorable node assignments, as unnecessary congestion has been removed in this region.
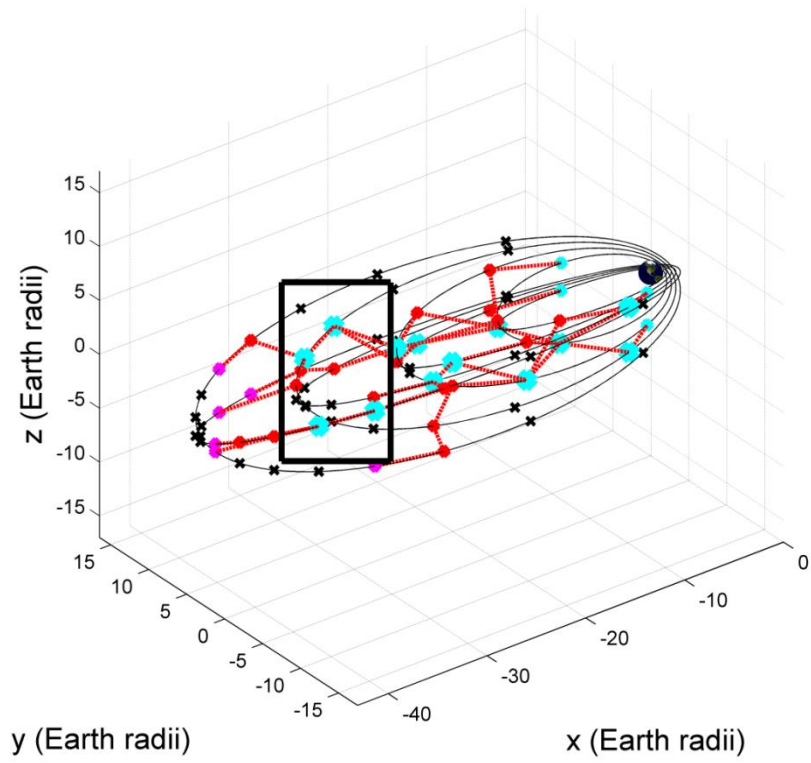
Fig. 12: Routes found without the anti-congestion mechanism. Turquoise nodes have more than one route passing through them, red nodes have only one.
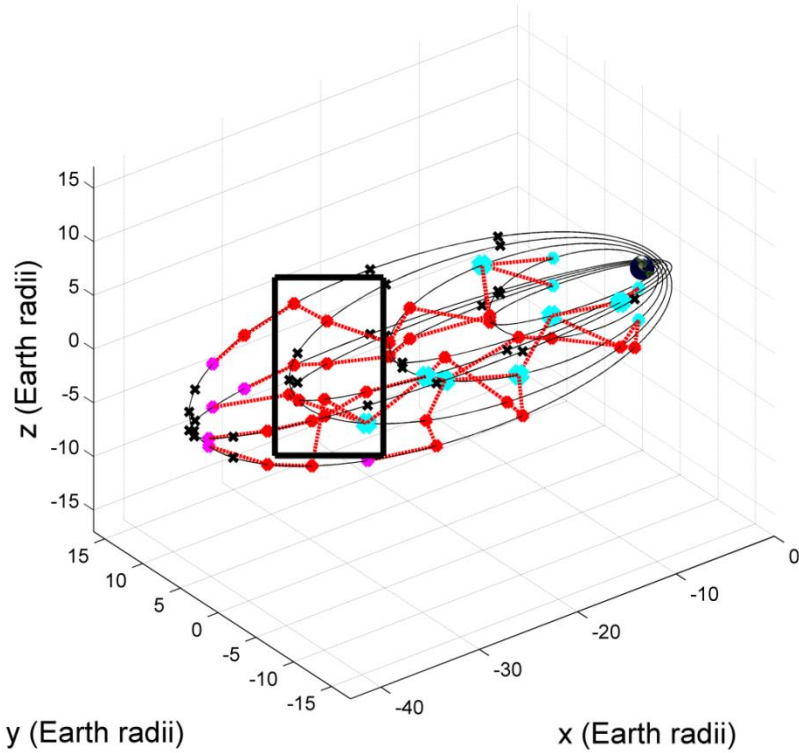

Fig. 13: Turquoise nodes have more than one route passing through them, red nodes have only one.

C. Real-Time Data Download

As a final benchmark of this algorithm's effectiveness, the amount of data downloaded to Earth during a single day is determined. The above method is applied to the test case from 48.5 hours to 72 hours after the given epoch (i.e. the third day after the initial time), with the search running every half-hour. The bandwidth is not simulated for the first two days to let satellites drift apart from their initial positions, so that any regularity in the network topology (due to the pre-set initial positions of the satellites) diminishes. It is assumed that the routing solution found at a time t is used until $t + 29$ minutes, after which a new search begins. This is essentially assuming that the satellites do not move significantly within half an hour. The total downloadable data in the time window is therefore the sum of the bandwidths achieved at each half-hour increment. Each satellite is assumed to be able to relay data at 10 Kbps, and each T_{sat} is transmitting scientific data at a rate of 5 Kbps. There are exactly 6 T_{sat} at every time step, near the apogee of the highest eccentricity orbit-pair. The total real-time data received at Earth using this method is 176.1 MB. This figure is to be related to the bandwidth of each satellite. Fig. 14 shows the bandwidth achieved at each time-step, along with the maximum downloadable data, which is determined by the number of R_{sat} at that time-step. These results show that this routing optimization method is viable and effective, as the bandwidth to Earth is always very close to the generated data rate. It is clear that no data is downloaded to Earth at the times where no satellite is in view of a ground station. To allow data to be downloaded at these times, changes would have to be made to the constellation itself. Additionally, the ratio of T_{sats} to R_{sats} is significant in determining how much bandwidth the constellation can provide; the scenario described in Section IV.B.1) has more R_{sat} bandwidth than scientific data being transmitted; however at other instants in time the total R_{sat} bandwidth can be lower than the data being transmitted, adding another source of loss for overall bandwidth.
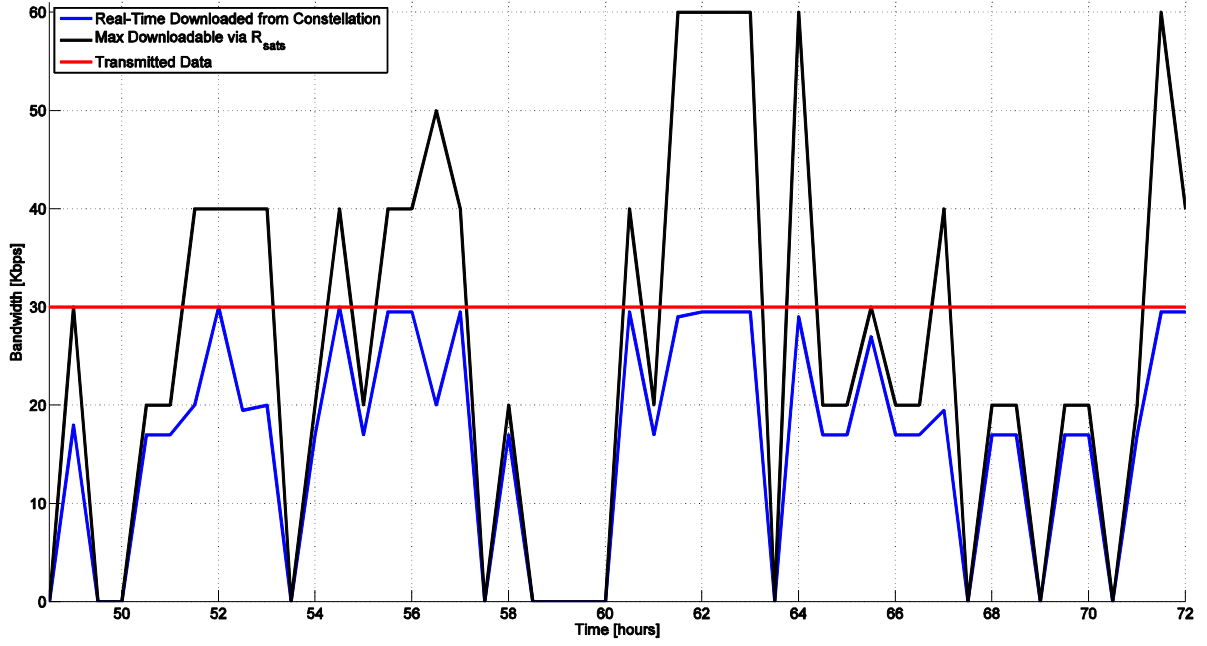
Fig. 14: Bandwidth achieved throughout one day.

V. CONCLUSIONS

The objective of this study was to design a network routing algorithm to maximize the real-time bandwidth to Earth in a constellation of inter-communicating satellites. The developed algorithm, based on Ant System (AS), is completely distributed, so a "central intelligence" is not necessary, instead each satellite, acting as node in the network, relies on the awareness of the availability of nearby satellites only. Such distributed system is advantageous as it is robust enough to dynamically and autonomously adapt to the changing positions of the satellites and the accessibility of the ground stations on Earth. The distributed system and the fact that routes are constantly being updated imply that isolated satellite failures are relatively inconsequential. Near-optimal link routes are found, being short (in terms of distance and/or number of relays), while also avoiding nodes with high load (congestion). The traditional AS algorithm had to be adjusted to suit this particular application; such as the inclusion of additional route prioritization to allow the algorithm to cope with the fact that multiple satellites are attempting to find routes to Earth and limiting ant lifetime to improve overall efficiency. In addition, tuning the parameters for the specific constellation plays an important role in maximizing the bandwidth.

Running this algorithm several times on one instant in the case study constellation, at the time $t = t_0 + 46$ hours and where the total amount of data to be downloaded to Earth is being transmitted at 30 Kbps, it consistently produced routes giving a bandwidth to Earth of 29.55 Kbps, with the difference being lost due to

data traffic congestion. A mechanism was put in place to reduce congestion during each route-finding run; it was shown that this prevented congestion where possible.

The results found using a simulated constellation of spacecraft (from mission study CENTINEL) proved that this is a viable way to establish real-time communications in a CubeSat mission for space weather monitoring. In this test-case scenario, where the T_{sats} transmit data at a rate of 5 Kbps and every other node in the constellation has a bandwidth of 10 Kbps, 176.1 MB of data per day were downloaded to Earth. It was found that this number could be significantly increased if more satellites with direct telecommunication link with Earth are available for longer. To increase the bandwidth to Earth of this constellation at any given time, more satellites should be placed on the inner two orbit-pairs.

Future work should consider how data that cannot be downloaded in real time due to congestion could be buffered, and downloaded at a later time. Further work should also study improvements in the anti-congestion mechanism.

With the CubeSat market increasing steadily, it is likely that large CubeSat constellations, possibly using high altitude and/or high eccentricity orbits will become widespread. This study provides a solution to the problem of real-time download to Earth for such missions.

VI. BIBLIOGRAPHY

[1] Chen, Z.; and Zeng, Y.; , "A Swarm Intelligence Networking Framework for Small Satellite Systems," *Communications and Network,* vol. 5, pp. 171-175, 2013. doi: 10.4236/cn.2013.53B2033

[2] Mortari, D.; , "Optimization of Flower Constellations for Dual Use," in *Aerospace Technologies and Applications for Dual Use*, River Publishers, 2008, pp. 105-113.

[3] "Iridium Homepage," Iridium Communications Inc., [Online]. Available: https://www.iridium.com/. [Accessed 2017].

[4] "Globalstar Homepage," Globalstar, Inc., [Online]. Available: https://www.globalstar.com/en/. [Accessed 2017].

[5] "OneWeb Homepage," OneWeb LLC, [Online]. Available: http://www.oneweb.world/. [Accessed 2017].

[6] Engelen, S; Gill, E; Verhoeven, C.; , "On the Reliability, Availability, and Throughput of Satellite Swarms," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 50, no. 2, 2014. doi: 10.1109/TAES.2014.120711

[7] Barnhart, D.J.; Vladimirova, T.; Sweeting, M.N.; , "Very-Small-Satellite Design for Distributed Space Missions," *Journal of Spacecraft and Rockets,* vol. 44, no. 6, 2007. doi: 10.2514/1.28678

[8] Eastwood, J.P.; , "Observing Magnetic Reconnection: The Influence of Jim Dungey," in *Magnetospheric Plasma Physics: The Impact of Jim Dungey's Research*, Springer International Publishing, 2015, pp. 198-199. doi: 10.1007/978-3-319-18359-6_9

[9] Pang, C.K.; Kumar, A.; Goh, C.H.; Le, C.V.; , "Nano-Satellite Swarm for SAR Applications: Design and Robust Scheduling," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 51, no. 2, 2015. doi: 10.1109/TAES.2014.140077

[10] Hadjitheodosiou, M.; Chen, Y.; , "A Case Study on Optimizing Communications for Constellation Space Missions," in *Wireless Communications and Networking*, New Orleans, 2003. doi: 10.1109/WCNC.2003.1200571

[11] Wood, L.; Pavlou, G.; Evans, B.; , "Effects on TCP of Routing Strategies in Satellite Constellations," *IEEE Communications Magazine,* vol. 39, no. 3, pp. 172-181, 2001. doi: 10.1109/35.910605

[12] Muri, P.; McNair, J.; , "A Survey of Communications Sub-Systems for Intersatellite Linked Systems and CubeSat Missions," *Journal of Communications,* vol. 7, no. 4, 2012. doi: 10.4304/jcm.7.4.290-308

[13] Bedon, H.; Negron, C.; Llantoy, J.; Nieto, C.M.; Asma, C.O.; , "Preliminary Internetworking Simulation of the QB50 Cubesat Constellation," in *IEEE Latin-American Conference on Communications*, Bogota, 2010. doi: 10.1109/LATINCOM.2010.5640977

[14] Donner, A.; Berioli, M.; Werner, M.; , "MPLS-Based Satellite Constellation Networks," *IEEE Journal on Selected*

Areas in Communications, vol. 22, no. 3, 2004. doi: 10.1109/JSAC.2004.823406

[15] Mushet, G.; Mingotti, G.; Colombo, C.; McInnes, C.; , "Self-Organising Satellite Constellation in Geostationary Earth Orbit," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 51, no. 2, 2015. doi: 10.1109/TAES.2014.130690

[16] Berioli, M.; Rosati, L.; Werner, M.; Reali, G.; , "Ant Routing Concepts for Dynamic Meshed Satellite Constellations," in *Advanced Satellite Mobile Systems Conference*, Noordwijk, 2004.

[17] Cavallaro, G.; Pham-Minh, D.; Bousquet, M.; , "HEO Constellation Design for Tactical Communications," in *IEEE AESS Conference on Satellite Telecommunications*, Rome, 2012. doi: 10.1109/ESTEL.2012.6400126

[18] Dorigo, M.; Stutzle, T.; , Ant Colony Optimization, Massachusetts Institute of Technology, 2004.

[19] Dorigo, M.; Maniezzo, V.; Colorni, A.; , "The Ant System: Optimization by a Colony of Cooperating Agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics,* vol. 26, no. 1, pp. 29-41, 1996. doi: 10.1109/3477.484436

[20] Gao, Z.; Guo, Q.; Wang, P.; , "An Adaptive Routing Based on an Improved Ant Colony Optimization in LEO Satellite Networks," in *Sixth International Conference on Machine Learning and Cybernetics*, Hong Kong, 2007. doi: 10.1109/ICMLC.2007.4370296

[21] Sigel, E.; Denby, B.; Hégarat-Mascle, S.; , "Application of Ant Colony Optimization to Adaptive Routing in a LEO Telecommunications Satellite Network," *Annales Des Télécommunications,* vol. 57, no. 5-6, pp. 520-539, 2002. doi: 10.1007/BF02995174

[22] Gao, K.; Wu, G.; Zhu, J.; , "Multi-Satellite Observation Scheduling Based on a Hybrid Ant Colony Optimization," in *2nd Internation Symposium on Computer, Communication, Control and Automation*, 2013. doi: 10.4028/www.scientific.net/AMR.765-767.532

[23] Zhang, Z.; Zhang, N.; Feng, Z.; , "Multi-Satellite Control Resource Scheduling Based on Ant Colony Optimization," *Expert Systems with Applications,* vol. 41, no. 6, pp. 2816-2823, 2014. doi: 10.1016/j.eswa.2013.10.014

[24] Battin, R.; An Introduction to the Mathematics and Methods of Astrodynamics, Reston, VA: American Institute of Aeronautics and Astronautics, 1999. doi: 10.2514/4.861543

[25] Kurose, J.F.; Ross, K.W.; , Computer Networking: A Top Down Approach Featuring the Internet, Pearson Education, 2003.

Neil Yarr received the M.Eng. degree in Aeronautical Engineering from the University of Glasgow (United Kingdom) in 2015 with a thesis on optimising inter-satellite telecommunications for real-time data download in CubeSat constellations.

Since 2015 he has worked as a Systems and Simulation Engineer at Clyde Space Ltd, Glasgow (United Kingdom) specialising in the development of CubeSat Attitude Determination and Control Systems.

Matteo Ceriotti received his M.Sc. summa cum laude from Politecnico di Milano (Italy) in 2006 with a thesis on planning and scheduling for planetary exploration. In 2010, he received his Ph.D. on "Global Optimisation of Multiple Gravity Assist Trajectories" from the Department of Aerospace Engineering of the University of Glasgow (United Kingdom). During 2009-2012, Matteo was a Research Fellow at the Advanced Space Concepts Laboratory, University of Strathclyde, Glasgow, leading the research theme "Orbital Dynamics of Large Gossamer Spacecraft". In 2012, he returned to the University of Glasgow as a Lecturer in Space Systems Engineering, within the School of Engineering. His main research interests are space mission analysis and trajectory design, orbital dynamics, trajectory optimisation, particularly focusing on high area-to-mass ratio structures.

Figure Captions

Fig.1: Ants have various route choices between source and destination; their decisions when travelling from node to node is probabilistic.

Fig.2: [8] Constellation of satellites, with the arc around the apogee of the outer orbit identifying the transmitting region. When a satellite enters this arc it changes its mode from data-relay node to T_{sat}, collecting and broadcasting scientific data.

Fig.3: (a) shows an illustration of a solution of the TSP: all nodes are visited and the start/end nodes can be anywhere. (b) is an illustration of the satellite telecommunications problem: the most optimal route must be found between two pre-defined nodes.

Fig.4: a), b) and c) illustrate route permissions 1, 2 and 3 respectively.

Fig. 5: Unnecessary congestion. Multiple T_{sats} using routes passing through a particular node, where other nodes are available.

Fig.6: The constellation of 74 satellites on eight nested orbits which comprise the CubeSat constellation proposal CENTINEL [8].

Fig.7: Spatial configuration of constellation for single-route test case analysis.

Fig.8: Algorithm progression with route permission 1. (a) number of ants which reach the R_{sat} without being terminated. (b) blue line: average length of routes; red line: length of the shortest route. (c) average number of hops taken by each ant in each generation.

Fig.9: Algorithm progression with route permission 3. (a) blue line: average length of routes, red line: length of the shortest route. (b) average number of hops taken by each ant in each generation.

Fig.10: An example optimal route. T_{sat} and R_{sat} are highlighted in blue.

Fig.11: Overall bandwidths achieved varying the weight.

Fig.12: Routes found without the anti-congestion mechanism. Turquoise nodes have more than one route passing through them, red nodes have only one.

Fig.13: Turquoise nodes have more than one route passing through them, red nodes have only one.

Fig.14: Bandwidth achieved throughout one day.