A Novel Approach to Quality-of-Service Provisioning in Trusted Relay Quantum Key Distribution Networks

Miralem Mehic, *Member, IEEE*, Peppino Fazio, *Member, IEEE*, Stefan Rass, *Member, IEEE*, Oliver Maurhart, *Member, IEEE*, Momtchil Peev, *Member, IEEE*, Andreas Poppe, *Member, IEEE*, Jan Rozhon, *Member, IEEE*, Marcin Niemiec, *Senior Member, IEEE*, and Miroslav Voznak, *Senior Member, IEEE*

***Abstract*— In recent years, noticeable progress has been made in the development of quantum equipment, reflected through the number of successful demonstrations of Quantum Key Distribution (QKD) technology. Although they showcase the great achievements of QKD, many practical difficulties still need to be resolved. Inspired by the significant similarity between mobile ad-hoc networks and QKD technology, we propose a novel quality of service (QoS) model including new metrics for determining the states of public and quantum channels as well as a comprehensive metric of the QKD link. We also propose a novel routing protocol to achieve high-level scalability and minimize consumption of cryptographic keys. Given the limited mobility of nodes in QKD networks, our routing protocol uses the geographical distance and calculated link states to determine the optimal route. It also benefits from a caching mechanism and detection of returning loops to provide effective forwarding while minimizing key consumption and achieving the desired utilization of network links. Simulation results are presented to demonstrate the validity and accuracy of the proposed solutions.**

***Index Terms*— Quantum key distribution, quality of service, routing protocol, real-time traffic.**

I. Introduction

DURING the 30 years since the discovery of the first quantum protocol [1], quantum technology has grown significantly and is rapidly approaching high levels of maturity. The next natural step in the evolution of quantum systems is to study their performance, suitability and convergence with applications used in everyday life. Significant progress in the development of quantum equipment has been reflected through a number of successful demonstrations of QKD networks [2]–[7] but without showing the clear suitability to assess how such networks compete with their classical counterparts under real-life conditions and in real-time traffic. The traffic in these networks was mainly considered with equal importance, and it was treated with the same priority. While such approach may be acceptable for some applications, it is not suitable for voice, video and collaborative applications. Since not all network traffic is equal, it should not be treated equally. Hence, different applications may have different service requirements with respect to the quality of service (QoS). This paper addresses the question of using real-time communication in QKD networks by considering QoS mechanisms. The primary goal is to provide an adequate QoS model that includes traffic classification and marking mechanisms, QKD link metrics that can be used to describe the state of the network accurately, and a scalable routing protocol that minimizes the consumption of key material through the equitable utilization of network resources.

QKD networks differ from traditional networks in several aspects:

- Although theoretical and pioneering results have been published in the field of quantum repeaters and quantum relays [8]–[10], in practice they remain unachievable with current technology[1] [11], [13]. Therefore, communication is realized in a hop-by-hop [2] or key relay manner [3], [14]. Both methods rely on the assumption that all nodes along the path between the sender and the receiver must be fully trusted, forming a trusted relay QKD network [2], [15], [16].

Manuscript received December 4, 2017; revised May 25, 2018 and December 15, 2018; accepted November 9, 2019; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor P. Papadimitratos. Date of publication December 17, 2019; date of current version February 14, 2020. This work was supported in part by the H2020 project OpenQKD under Grant 857156, in part by the institutional grant SGS conducted at the Technical University of Ostrava under Grant SP2019/41, in part by the ESF in Science without borders project within the Operational Programme Research, Development and Education under Grant CZ.02.2.69/0.0/0.0/16 027/0008463, and in part by the Ministry of Education, Science and Youth of Canton Sarajevo, Bosnia and Herzegovina, under Grant 11/05-14-27719-1/19. *(Corresponding author: Miralem Mehic.)*

M. Mehic is with the Department of Telecommunications, Faculty of Electrical Engineering, University of Sarajevo, 71000 Sarajevo, Bosnia and Herzegovina, and also with the VSB–Technical University of Ostrava, 708 33 Ostrava-Poruba, Czech Republic (e-mail: miralem.mehic@ieee.org).

P. Fazio, J. Rozhon, and M. Voznak are with the VSB–Technical University of Ostrava, 708 00 Ostrava, Czech Republic.

S. Rass is with the System Security Group, Institute of Applied Informatics, Universitaet Klagenfurt, 9020 Klagenfurt, Austria.

O. Maurhart and A. Poppe are with the Security and Communication Technologies, Center for Digital Safety and Security, AIT Austrian Institute of Technology GmbH, 1220 Vienna, Austria.

M. Peev is with the Optical and Quantum Laboratory, Munich Research Center, Huawei Technologies Duesseldorf GmbH, 80992 Munich, Germany.

M. Niemiec is with the Department of Telecommunications, AGH University of Science and Technology, 30-059 Krakow, Poland.

Digital Object Identifier 10.1109/TNET.2019.2956079

[1]The idea behind implementing quantum routers is to use quantum entanglement of photons to communicate over different quantum channels. In short, it means that multiple particles are linked together such that the measurement of one particle's quantum state determines the possible quantum states of the other particles. Even when the particles are separated by a considerable distance, they still make up a joint quantum system. The fidelity of a quantum state decreases exponentially with the distance of its qubits due to noisy quantum channels [11], [12]. In addition, quantum memory is required to implement a quantum repeater according to [9], [10]. Although implementations of quantum memories exist, which can store a qubit for between several milliseconds to one second or even more, this is still too short for practical applications.

- Nodes are connected with QKD logical links, referred to below as links, which employ two distinct channels: a quantum channel, which is used for transmission of raw cryptographic keys encoded in certain photon properties, and a public channel, used for verification and processing of the exchanged values. Each quantum channel is always a point-to-point connection between exactly two nodes [17], while public channels can be implemented as any conventional connection which can include an arbitrary number of intermediate devices [18].
- The key rate is interconnected with a length of optical fiber such that a longer distance implies a lower key rate due to absorption and scattering of photons [11], [19]–[22]. Although key rates of 1 Mbps and above have been achieved [5], [23]–[25], such solutions are limited to very short distances. Therefore, both endpoints of the corresponding link implement key buffers (storages) of limited capacity, which are gradually filled at their maximum key rate with the processed cryptographic key, referred to below as key material, and subsequently used for encryption/decryption of data flow [17], [26]. Key material denotes the symmetric cryptographic keys that are generated during the QKD process, stored in buffers (storages) and used subsequently for cryptographic operations over user traffic.
- Without key material, cryptographic operations cannot be performed, and a link can be described as temporarily unavailable [15]. To provide information-theoretically secure (ITS) communication, the key tends to be applied with a One-Time Pad (OTP) cipher and authenticated using an ITS message authentication scheme such as Wegman-Carter when communicating over a public channel [27]–[29]. As a result, ITS communication requires more bits of key material than the length of the secured message [17]. The type of encryption algorithm used and the volume of network traffic to be encrypted determines the key storage emptying speed, referred to as the key consumption rate. The key consumption rate denotes the rate of key material being taken from buffers (storages) and used for cryptographic operations. Similarly, the charging key rate (or simply key rate) denotes the rate of new key material generation, that is, the rate of adding new keys to the buffers (storages) [3], [17], [30].
- To meet the requirement to bypass untrusted nodes, in practice, QKD networks are usually deployed as overlay point-to-point networks which exhibit selfish behavior, acting strategically to optimize performance and resulting in dynamic and unpredictable link performance [17], [26], [31]–[33]. Overlay networks use existing underlying networks in an attempt to implement a better service; one of their most important features is the independence of the path offered by Internet service providers (ISP).

This paper is organized as follows: Section II covers related work, Section III points out the significant resemblance between QKD and ad-hoc networks, and Section IV defines the requirements for providing QoS in QKD networks. In Section V we propose a novel QoS model including a GPSRQ routing protocol. The simulation setup is presented in Section VII, with Section VII-A presenting the evaluation of obtained results. Section VIII concludes the paper.

II. Related Work

Given the common assumption that all nodes along a path in a QKD network must be fully trusted [15], [16], QKD networks were mainly analyzed from two aspects: security and network performance. The idea of passive eavesdropping, in which the adversary may use eavesdropping not to extract information but to redirect the data flow towards a node under their control, has been analyzed in [34]. Following a similar idea, stochastic routing has been proposed to avoid deterministic routing which is used in traditional routing protocols [35], [36]. Game-theoretic techniques have been used to find an optimal balance between interdependent service quality criteria with distinct performance indicators [37].

QoS in QKD networks deployed previously has been largely neglected, stating that it is somehow achievable without any difficulties. Therefore, the Open Shortest Path First (OSPFv2) routing protocol has been modified to use the amount of key material as the routing metric while ignoring the performance of the public channel [26], [38]. In [39], the author proposes using unencrypted and non-authenticated communication for the dissemination of OSPFv2 routing packets; this is simple prey for an eavesdropper who is assumed to have unlimited resources at their disposal, especially when passive eavesdropping is taken into account [34]. In [40], [41], unmodified OSPFv2 was combined with reserving key material resources. Considering the interdependence of the public and quantum channels [42], reservation of key material resources does not solve the problem of QoS since the routing path may be inadequate for quantum channels.

III. Similarities Between QKD and MANET Technologies

The specific QKD issues and constraints described above pose significant challenges in QKD network design. However, by analyzing the characteristics of QKD networks, we note similarities with Mobile Ad Hoc Networks (MANET) [43]–[45]. First, we specify the main characteristics of QKD technology from a simple point of view:

- QKD links, described above, due to features of quantum channels, are always implemented in point-to-point behavior, and they can be roughly characterized by two features: limited distance and key rate inversely proportional to the distance [17]. Additionally, QKD links may become unavailable when there is no enough key material or when the public channel is congested [42]. Such behavior is similar to Wi-Fi links which are limited in length and where the communication speed depends on the user's distance from the transmitter.
- One of the main features of QKD networks is the absence of a quantum repeater or quantum router in practice, therefore communication is usually performed on a hop-by-hop basis [13], [17].

In MANET, communication takes place on a hop-by-hop basis, and mobile nodes are typically powered by batteries,

placing special attention on energy-aware solutions. The nodes connect themselves in a decentralized, self-organizing manner with no authority in charge of managing and controlling the network [44], [46]. The battery power in MANET nodes can be easily linked to the amount of key material in QKD key storages. Given the node without a power supply (empty batteries) is not an active member of the network, the same analogy is valid for QKD networks where the node without available key material cannot be used for data transmission. The range limitations of wireless links can be mapped to the limitations in the length of QKD links, while the absence of dedicated network infrastructure (such as a router) is common to MANET and QKD. Here we recognize the significant similarity between these two technologies, allowing us to propose a new approach to addressing QKD network issues.

IV. QoS in QKD Networks

The specific QKD constraints described above lead to the conclusion that this type of network provides weak support to QoS.

A. QoS Models

1) Integrated Service and QKD Networks: The basic concept of the Integrated Service (IntServ) model is a per-flow resource reservation using the Signaling Resource ReSerVation Protocol (RSVP) before data transmission [47]. In our view, the IntServ model is not suitable for QKD networks because of the following:

- *An inability to guarantee the reservation:* when QoS is provided in an IP network, an IP router has complete control over its packet buffers and the output link bandwidth, and can directly schedule these resources. In contrast, in an overlay network, the node cannot directly access the available resources in the overlay path. It can only rely on measurement techniques where high accuracy cannot be guaranteed and it cannot directly control or reserve resources in the underlying network. The only thing that a node in a QKD network is able to do is to guarantee resources of a quantum channel by reserving key material in key storages. However, considering the interdependence of public and quantum channels in a QKD link [42], such reservation does not constitute any gain.
- *Signaling:* RSVP is an out-of-band signalization protocol which means that signaling packets contend for network resources with data packets and consume a substantial amount of scarce key material and network resources.

2) Differentiated Service and QKD Networks: Differentiated Service (DiffServ) uses Differentiated Services Code Point (DSCP) bits in the IP header and a base set of packet forwarding rules known as Per-Hop-Behavior. DiffServ is known as an edge-provisioning model, and it does not provide any QoS guarantees *per se* [48]. The application of DiffServ in its original form is limited in a QKD network due to the following:

- *Edge router selection*: existing QKD technology limits the deployment of a QKD network to the metropolitan scale[2] [2], [3], [5], [6], [49], [50]. In such a network, it is necessary to clearly define the edge routers which play a key role in the processing of traffic.
- *Lack of Service Level Agreement*: each network node needs to comply with the rules for the classification and processing of traffic of different priorities. However, since the Service Level Agreement (SLA) concept is not defined in a QKD network, it is questionable how nodes of potentially different domains can negotiate traffic rules.

B. QoS Signalization in QKD Networks

In general, the QKD protocol which establishes a new key material consists of six successive stages: a physical exchange of quantum states between a pair of devices, extraction of the raw key (sifting), error rate estimation, reconciliation, privacy amplification, and authentication [51], [52]. Only the first stage is performed over a quantum channel; all other stages are performed over a public channel, resulting in the communication referred to as QKD post-processing. We propose an extension of authenticated packets with signaling data which provides an elegant way of tackling the problem of distribution of signaling information without introducing additional traffic overheads [53].

C. QoS Routing in QKD Networks

In our view, a routing protocol well-suited for operation in dynamic QKD networks should meet the following main design objectives:

- Reducing the consumption of scarce key material by choosing the shortest path considering both channels of the QKD link. The routing algorithm needs to find a balance between the requirements, since a path that meets the requirements of the public channel may not be suitable for the quantum channel and vice versa [17], [42].
- Given that the main objective of QKD is to provide ITS communication, routing packets need to be encrypted and authenticated [17], [54]. This means that the number of routing packets needs to be minimized to preserve scarce key material.
- To prevent denial of service, it is necessary to minimize knowledge about the utilized routing path by reducing the broadcast of routing packets [34].
- The routing protocol should be scalable to different network sizes.
- Due to a low key charging rate and overlay networking mode, link interruptions are common in QKD networks. Hence, the routing protocol should be robust enough to find an adequate replacement path.

In general, routing solutions can be divided into three broad categories: source, hierarchical and distributed routing. The performance of source routing algorithms relies on the availability of precise link state information, while the dynamic nature of QKD networks makes the available

[2]In 2014, a QKD system connecting the cities of Hefei-Chaohu-Wuhu (HCW) in China with a total of nine QKD nodes was reported in [6]. Thirteen QKD devices within nine nodes in total were employed to support the two QKD networks and the intercity QKD link.

link state information inherently imprecise. Given that the constant maintenance of link state information is mostly done by periodic flooding, this solution is inadequate for QKD networks. Although several examples exist of hierarchical network organization [4]–[6], in our opinion, such organization is not suitable for QKD networks since nodes of upper hierarchical levels represent a potentially easy target for the attack to disassemble the network. In distributed routing, the computation of the path is shared among network nodes on a periodic basis (proactive) or only when a routing path is requested (reactive). Proactive routing protocols mainly use the static update period time for keeping routes up-to-date, which contradicts the dynamic nature of QKD networks. Therefore, in overlay networks, reactive routing performs better in terms of efficiency and stability than proactive routing [55], [56].

V. FQKD: A Flexible Quality of Service Model for QKD Networks

To overcome problems of providing QoS in a dynamic environment, we present a flexible QoS model for QKD networks (FQKD). Our model avoids a centralized resource management scheme or reservation of resources mechanism. Instead, we turn to a distributed approach to control traffic loads by providing soft-QoS constraints without flow or session state information maintained in support of end-to-end communication. FQKD defines three roles for nodes in a QKD network: ingress, interior and egress. Each node can take any of these roles depending on the position in network flow. A source node which sends data is referred to as an ingress node. Interior nodes are nodes that forward data to the final destination node, which is referred to as an egress node.

A. Provisioning and Conditioning

As shown in Fig 1, the FQKD model consists of a sender-based classifier, waiting queues, local node-based admission controller, crypto module and dynamic regulation of admitted sessions at the MAC layer. The classifier is input at the ingress node to distinguish between traffic classes by marking the DSCP field in the IP packet header. FQKD distinguishes between three traffic classes with corresponding DSCP values: best-effort, real-time and premium class. For each class, separate waiting queues are defined and processed by priority. The packets are forwarded by interior nodes in per-hop behaviour associated with the assigned DSCP value.

Considering that nodes in QKD networks continuously generate new keys at their maximum rate[3] [26], before setting the route, the routing protocol contacts the admission controller to filter those links to its neighbors that have sufficient resources to serve the classified network packet. The routing protocol calculates the path and the packet is forwarded to the MAC layer for further processing. Otherwise, if no available link is found, the packet waits in the queue for reprocessing. In FQKD, additional waiting queues are installed between the

[3]QKD nodes aim to generate as much key material as possible. Therefore, nodes continuously generate traffic at the highest rate allowed by the network. More information about the impact of public channel states on the intensity of the key rate can be found in [42].

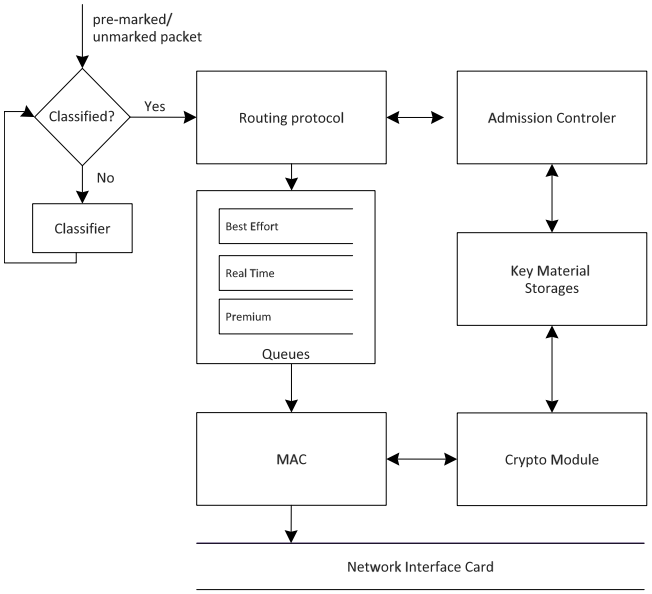

Fig. 1. FQKD model. First, the incoming application package is classified and forwarded to the routing protocol which contacts the admission control to verify there are sufficient resources for processing the data. If true, the packet is stored in the appropriate waiting queue and encrypted before sending to the network. Arrows denote communication between model elements and the flow of the processed packet.

L3 and L4 TCP/IP layer to avoid conflicts in decision making which could lead to inaccurate routing. Suppose the queues are implemented on the data link layer (L2) only, and suppose that they are half filled with packets. Since the routing protocol used the routing metric that at the time t_1 of calculating the route had a different value from the time t_2 when the packet came online in the queue to be served, it follows that significant changes to the state of links in the time interval $\Delta t = t_2 - t_1$ could occur, which can lead to inaccurate and incorrect routing. Instead, by exploiting higher level waiting queues, the packet for which the route is calculated will be directly forwarded to lower layers and immediately sent to the network. This implies the usage of one set of waiting queues (set of three waiting queues for best-effort, real-time and premium traffic classes) for all network interfaces. Using queuing at the L2 layer is not excluded, but additional attention is given to queues at a higher level due to the dynamic nature of the network.

Assuming the key rate is constant in time when the quantum channel has a fixed length [17], [26], it is evident that the key storage can be identified with the Token Bucket traffic shaping mechanism as shown in Fig. 2. This simplifies the view of the admission controller, which behaves as the traffic conditioner. The volume of traffic over the QKD link is limited by the amount of key material in key storage which is used for encryption or authentication of data over that link. The key material storage of link k between nodes a and b can be represented using following parameters:

- The time measurement moment t,
- The average key generation rate r_k, measured in bits per seconds and used to indicate the charging rate of the storage,

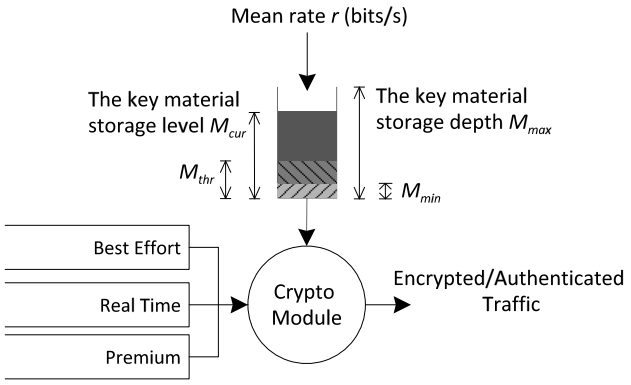

Fig. 2. Traffic processing in the FQKD model.

- The key material storage depth $M_{max,k}$, used to indicate the capacity of the storage,
- The current value $M_{cur,k}(t)$, representing the amount of key material in the storage at the time of measurement *t*, where it holds that $M_{cur,k}(t) \leq M_{max,k}$.
- The threshold value $M_{thr,a,b}(t)$ or simply $M_{thr,k}(t)$ described in more detail below,
- QKD is also known as Quantum Key Growing [57], [58] or Quantum Key Expansion [59], since it needs a small amount of key material pre-shared between parties (denoted with $M_{min,k}$) to establish a larger amount of the secret key material. This pre-shared key material is used for QKD post-processing and authentication [29], [59].

The amount of key material over a time interval T depends on the key generation rate and the available amount of key material in the storage, and can be calculated using Equation (1). The average operational rate, that is the rate at which packets can be served over an interval T, can be calculated using Equation (2). As $T \rightarrow \infty$, the operational rate $A_k(T)$ approaches the charging key rate r_k.

$$D_k(T) \leq r_k \cdot T + M_{cur,k}(T) - M_{min,k}(T) \quad (1)$$

$$A_k(T) = \frac{D_k(T)}{T} = r_k + \frac{M_{cur,k}(T) - M_{min,k}(T)}{T} \quad (2)$$

However, the overall amount of traffic data that can be transmitted over link k in a time interval T can be calculated by dividing the value obtained from Equation (1) and the ratio L_k which is the quotient of key length used for encryption with authentication and the length of the data message. Note that for ITS communication, which usually involves the OTP cipher and Wegman-Carter authentication, more bits of key material than the length of the data message are required ($L_k > 1$) [17].

An incoming packet is served from the queue only if there is enough key material in the storage. Otherwise, the packet remains in the queue waiting for storage to be charged. The length of the queue is limited, and traffic-shaping algorithms are in charge of packet management operations. To avoid blocking work due to the lack of the key material used to generate new key material, special attention is placed on the categorization of the traffic. If storage stops charging, the purpose of the link loses functionality. Therefore, the traffic generated by the post-processing application has the highest priority and is sorted in the premium queue. Only traffic from post-processing applications can use the key material when $M_{cur,k}(t) \leq M_{min,k}$ while the traffic from the other two queues is served only when $M_{cur,k}(t) > M_{min,k}$.

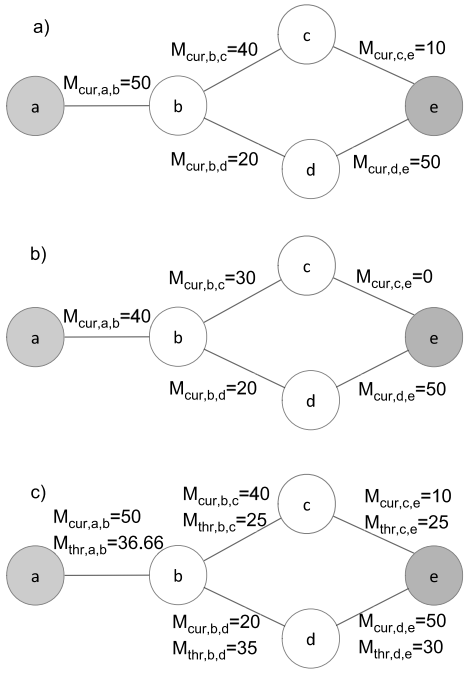

Fig. 3. Simple topology showing the calculation of M_{thr}: a) the traffic is routed along the route *a-b-c-e*; b) the traffic is routed along the path *a-b-c* regardless of key material depletion of link *c-e*; c) calculated M_{thr} values based on M_{cur} values of all links connected to a given node.

The threshold value M_{thr} is proposed to increase the stability of QKD links, where $M_{thr,k}(t) \leq M_{max,k}$. The parameter is explained by considering the simple topology shown in Fig. 3 where node *a* needs to communicate with remote node *e*. Suppose the routing protocol uses only information about the state of its links with its neighbors. Then, assuming that all network links have the same performance of public channels, we consider only the state of key storages which are marked next to links as shown in Fig. 3-a. Upon receipt of the packet from node *a*, the routing protocol on node *b* selects path *b-c* since the link *b-d* has a lower performance. However, if node *b* does not consider the state of links that are more than one hop away, the traffic may stuck on the link between nodes *b* and *c* as shown in Fig. 3-b. To avoid such behaviour, we propose using the M_{thr} value. Each node *i* calculates value L_i summarizing the M_{cur} values of links to its neighbors and dividing it by the number of its neighbors N_i using Equation (3). Then, each node exchanges calculated value of L_i with its neighbors. The minimum value denoted by Equation (4) is accepted as a reference threshold value of the link.

$$L_i = \frac{\sum_{k=0}^{N_i} M_{cur,i,k}}{N_i} \quad (3)$$

$$M_{thr,i,j} = min\{L_i, L_j\} \quad (4)$$

As shown in Fig. 3-c, node *b* calculates $L_b = 36.66$, while node *c* calculates $L_c = 25$. The threshold value of the link

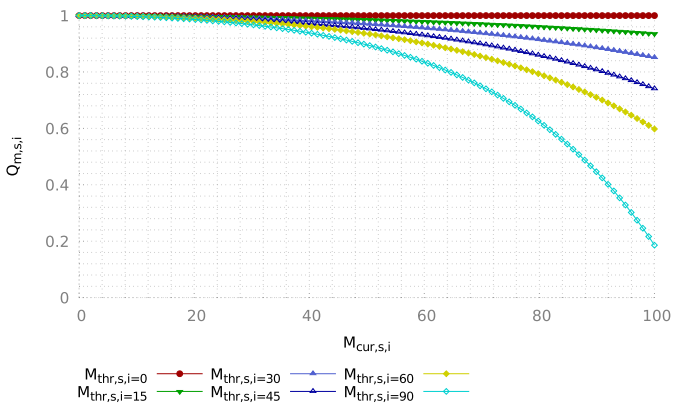

Fig. 4. Q_m of link between nodes s and i for different values of M_{thr}; $M_{max,s,i} = 100$.

b-c is set to $M_{thr,b,c} = 25$ and it is included in link metric calculation as described in Section V-B. The higher the value of M_{thr}, the better the state of links that are more than one hop away.

B. QKD Link Metric

Popular metrics from conventional networks which describe the state of the communication link cannot be adequately used in QKD networks since they only describe the public channel. Therefore, we propose new metrics that clearly define the state of the QKD link taking into account its most essential features.

1) The Quantum Channel Status Metric: At the moment of serving the packet, the remaining key material in the key storage is the main factor contributing to the link's availability; this is because without key material, cryptographic operations cannot be performed and secure communication over the link is not possible. We use Equation (6) to express the state of the quantum channel between nodes *s* and *i*, where $Q_{frac,s,i}$ is the ratio of the squared amount of key material at the time of measurement (${M_{cur,s,i}}^2$) multiplied by the threshold value ($M_{thr,s,i}$) and the cubed capacity of the key storage (${M_{max,s,i}}^3$) as defined by Equation (5). $Q_{frac,s,i}$ is in the range [0,1], and it highlights the current amount of key material on closest links in relation to the amount of key material of links that are further away since more distant links are unreachable when links to neighbors are unavailable.

$$Q_{frac,s,i} = \frac{{M_{cur,s,i}}^2 \cdot M_{thr,s,i}}{{M_{max,s,i}}^3} \tag{5}$$

$$Q_{m,s,i} = 1 - \frac{Q_{frac,s,i}}{e^{(1-Q_{frac,s,i})}} \tag{6}$$

$Q_{m,s,i}$ is the utility function associated with the key material level of the link. $Q_{m,s,i}$ uses an exponential formula to address the fact that the lower the amount of key material in the storage, the more critical the situation, and the less time is left for the routing protocol to react.

The value of $Q_{m,s,i}$ is normalized as a grade ranging from 0 to 1 as shown in Fig. 4, where a lower value means a better quantum channel state. In the example shown in Fig. 3-c, the routing protocol should favor the link *b-d* since $Q_{m,b,d} < Q_{m,b,c}$.

2) The Public Channel Status Metric: Instead of using popular approaches from conventional or overlay networks [31], [60], we use meta-data of keys such as the time duration of the key establishment process to assess the state of the public channel effectively. We define $P_{m,s,i}$ with Equation (7) to evaluate the state of the public channel between nodes *s* and *i* where $T_{last,s,i}$ is the length of time spent on the establishment of the key material at the time of measurement and $T_{maximal,s,i}$ is the maximum time that can be tolerated for the establishment of the key. $T_{maximal,s,i}$ is calculated as the double value of the average duration of key material establishment process, in the long run, denoted as $T_{average}$ in Equation (8).

$$P_{m,s,i} = \frac{T_{last,s,i} + \Delta t}{T_{maximal,s,i}} \tag{7}$$

$$T_{maximal,s,i} = 2 \cdot T_{average} \tag{8}$$

Δt is used to describe the freshness of the information and is defined as the difference between the current time of the measurement and the time when the $T_{last,s,i}$ is recorded. Note that $T_{average}$ is not equal for all links of the network since it depends on the load of the network, types of quantum and network devices, QKD post-processing application and the state of the public channel.[4] The value of $P_{m,s,i}$ mainly falls in the range [0,1] where the lower value means a better public channel state. Values greater than 1 indicate that the link has a problem with establishing new key material, and such links should not be considered.

3) The Overall QKD Link Status Metric: Key material depletion is not the same for all links since it depends on the type of encryption algorithm used and the volume of network traffic to be encrypted. For example, a QKD link between nodes *s* and *i* having a low value of $Q_{m,s,i}$ may be suitable for a network flow encrypted with less secure algorithms which require less key material than OTP (such as the AES cipher), but it may not be suitable for encryption using OTP. Thus, factor α which reflects the balance between the requirements is introduced in Equation (9) to compensate for this effect using the utility functions of the quantum and public channel in the [0,1] range where a lower value means a better overall link state.

$$R_{m,s,i} = \alpha \cdot Q_{m,s,i} + (1 - \alpha) \cdot P_{m,s,i} \tag{9}$$

Parameter α takes the value in the [0,1] range; if the OTP cipher is used, we suggest the value of $\alpha = 0.5$. This means that both channels of the QKD link are considered equally. If the AES cipher is used, α can be set to a lower value to

[4]In practice, the QKD link employs three channels. The QKD quantum-channel and the QKD distillation-channel are usually referred to as the quantum channel and public channel, respectively. The third channel is known as the (time-stable) synchronization channel and cannot be separated from the quantum channel physically to two fibers, because it delivers time references. From the perspective of higher network layers, the synchronization channel is part of the key generation (growing) process. If the impact of this channel on the entire process were to be considered, the signals in the synchronization channel need to be sufficiently large so that even co-propagating classical channels or spoiled signals are irrelevant and key generation is disturbed. Such major non-synchronization would be reflected in the overall duration of the key generation process, which is considered using Equation (7).

put a greater emphasis on the public channel due to its lower requirements for key material, depending how frequently it is refreshed and the length of the AES key used.

VI. Greedy Perimeter Stateless Routing Protocol for QKD Network

Driven by the similarities between the MANET and QKD networks discussed in Section III, we present the Greedy Perimeter Stateless Routing Protocol for QKD networks (GPSRQ). The primary motivation for designing GPSRQ is to minimize the number of routing packets and to achieve high-level scalability by using distributed geography reactive routing.

We assume that all nodes know the geographical locations of all other network nodes they wish to communicate with. Therefore, there is no periodic flooding of the node location details, and we assume a location registration and lookup service that maps the node address to a location. This paper does not deal with implementation details of such a service, but we assume it can be implemented using internal or other communication channels [61]. As indicated in Section IV-B, authenticated packets in QKD post-processing can be used to exchange information about the geographical position of nodes effectively. Although several experiments have been conducted regarding mobile QKD networking [62]–[66], due to the high sensitivity of quantum equipment to various environmental factors[5] and relatively low key generation rates in free-space QKD links [24], [59], [68]–[70], in this paper we assume that QKD networks are composed of static nodes representing secure access points[6]. We follow the idea "*the greater the distance separating two nodes, the slower they appear to be moving with respect to each other*" outlined in [71] to implement caching in GPSRQ, which is discussed further below.

GPSRQ sets up a network without hierarchical organization, which means that all nodes in the network are of equal importance. Nodes do not exchange routing tables, which significantly minimizes the consumption of scarce key material and reduces the probability of passive eavesdropping [34]. Route selection, that is the decision about the next hop, is made in per-hop behavior such that the packet is moved closer to the destination based on the states of links in the local environment and on the geographical distance from the node. An eavesdropper is not able to intercept routing packets and find out the exact route to the destination, since it is not known at which node's network interface the packet will be forwarded until the last moment. GPSRQ uses two packet forwarding algorithms: greedy forwarding and recovery mode forwarding.

A. Greedy Forwarding

By definition, greedy forwarding entails forwarding to the neighbor geographically closest to the destination. An example of greedy forwarding is shown in Fig. 5, where ingress node a which is surrounded by three adjacent nodes b, d, and k, needs to communicate with egress node g. Ingress node a forwards the packet to b, as the Euclidean distance between b and g is smaller than the distance between g and any of a's other neighbors. Such greedy forwarding is repeated on interior nodes and it stops when the packet reaches its destination. GPSRQ aims to maximize network utilization by using different paths for different traffic classes. It uses Equation (10) to calculate the path to forward the packet:

$$F_{a,g,v} = \big(1-\beta\big)\cdot R_{a,v} + \beta \cdot G_{v,d} \tag{10}$$

where $R_{a,v}$ denotes the state of link between source node a and neighboring node v using Equation (9) and $G_{v,g}$ represents the Euclidean distance between neighboring node v and destination node g, for each node v which belongs to the set N_a of all neighbors of source node a, $\forall v \in N_a$. All routes towards the destination are sorted in descending order using Equation (10) and the route with the lowest value is used. GPSRQ uses parameter β with a value in the [0,1] range to manage network utilization by choosing between forwarding along the geographically shortest route or the route that has the most available resources. We discuss simulation results for different values of β in Section VII-A5.

Greedy forwarding relies on the knowledge of the geographical location and state of links to neighbors, which results in a high level of network scalability. However, in some cases, the route to the destination requires forwarding the packet over a neighboring node which is geographically further from the destination than the node which forwards the packet. In cases when a local maximum occurs, an alternative recovery-mode forwarding is used. To increase scalability and exclude routes which do not lead to the destination, we propose a robust caching mechanism to preserve key material consumption. Consider the example shown in Fig. 5 where we further suppose that link *b-f* is unavailable due to a lack of key material. When node b realizes there is only one interface

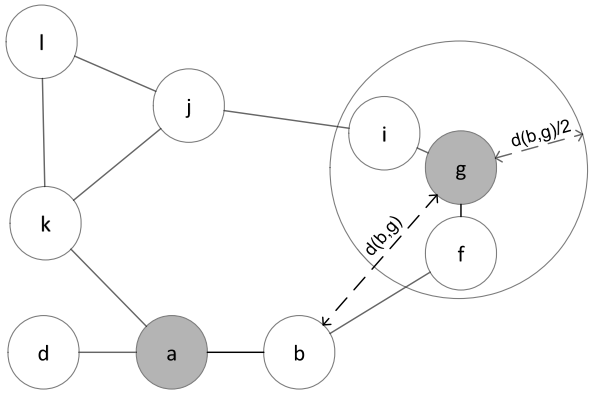

Fig. 5. Ingress node *a* which is surrounded by three adjacent nodes *b*, *d*, and *k*, aims to communicate with the egress node *g*. In the absence of a path over node *b*, node *a* writes in the internal cache that it is not possible to route packets along the path *a-b* towards the region marked with a circle of radius *d(b,g)/2* with the center in node *g*. Any further request for routing towards any node which is placed in the defined circle region will be ignored over the route via node *b*, and an alternative route needs to be found.

[5]The key rate may vary due to humidity, temperature, stability of devices, global radiation, pressure, dust, sunshine duration or other factors [17], [67].

[6]Low mobility within a geographic region or slow node movement are supported as long as the service for distributing geographic locations of nodes can accurately and safely distribute locations to all nodes in the network.

available (the interface which was used to receive the packet from node *a*), it marks a "loop" field in the GPSRQ packet header and returns the packet back to node *a*. Then, node *a* calculates the Euclidean distance *d(b,g)* between node *b* and destination node *g* and writes in its internal cache memory that it is not possible to route towards the region which is marked with a circle of radius *d(b,g)/2* and the center in node *g* along the path *a-b*. Upon receiving any further requests for routing towards the node placed in the defined circle region, node *a* will ignore the route over node *b* and look for an alternative route. The validity of cached record is set to time interval defined using Equation (11):

$$T_{cache} = T_{maximal,b,g}/2 \tag{11}$$

where $T_{maximal,b,g}$ is defined by Equation (8). The overall goal is to reduce the dynamics of the network topology changes using the scalable cache mechanism. In Section VII-A6, we discuss the impact of different values on cache validity. After the cached record expires, the node is allowed to try establishing the connection once again. In the event when GPSRQ detects there is no neighbor closer to the destination, it enters recovery mode.

B. Recovery Mode Forwarding

Recovery mode involves using the well-known right-hand rule which states that the next edge from node *k* upon arriving from node *a* is edge *(k,j)* which is sequentially counterclockwise to edge *(a,k)* [72]. However, for first forwarding, the packet is forwarded along edge *(a,k)* which is counterclockwise to node *a* from line $\overline{ag}$. The packet stays in recovery mode until it reaches a node which is closer to the destination than the node where forwarding in recovery mode started. To avoid routing loops, the GPSRQ header contains information about the IP address of the node at which the packet entered recovery mode and the outgoing interface which was used for first forwarding. If the loop is detected by analyzing the packet header, the packet is returned to the previous node for rerouting and adding a new entry to the node's internal cache memory. If the public channel of link *j-i* is unavailable, there is no available path to destination *g*. Source node *a* forwards the packet to node *k* which forwards the packet to node *j* in greedy forwarding mode. Since link *j-i* is unavailable, node *j* is not able to find any neighbor closer to the destination so it enters recovery mode, sets the value of the field "inRec" to 1, writes its IP address to the header field "recPosition", writes the interface number which leads to node *l* in the "recIF" header field and forwards the packet to node *l* since it is on the first edge counterclockwise about *j* from the line $\overline{jg}$ as required by the right-hand rule. Upon receipt of the packet, node *l* inspects the header and remains in recovery mode since node *j* in which the packet entered recovery mode is closer to the destination *g*. Then, the packet is forwarded to node *k* which forwards the packet back to node *j*. When node *j* detects its IP address from the header field "recPosition", it adds a record to the internal cache memory stating that it is not possible to reach destination node *g* via node *l* by adding a record consisting of the triple: IP address of hop *l*, the radius of circle region *d(l,g)/2* and the value of the circle center which is set to the location of node *g*. Given that there is no other interface available, node *j* sets GPSRQ header field "loop" to 1 and returns the packet to node *k* which adds a record to its cache memory stating that it is not possible to reach destination node *g* via node *j*. Then, node *k* sets the field "loop" to 2 and tries again with greedy forwarding where node *j* is excluded as the next hop. The packet is forwarded to node *l* which forwards the packet to node *j*. Since node *j* does not have any other interface available, it will set the value of header field "loop" to 1, and return the packet to node *l* which will add to its cache memory a record stating that it is not possible to reach destination node *g* over node *j*. The packet is returned to node *k* and then it is returned to node *a*. This procedure is repeated until a feasible path to the destination is found. Otherwise, if no available path is found, the packet is gradually returned to the source node which is allowed to discard the packet while keeping updated records in the cache memory of other nodes.

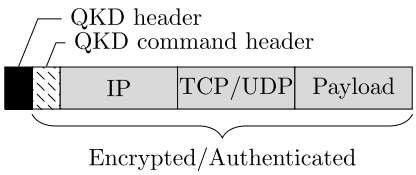

Fig. 6. Packet encapsulation with the QKD header and QKD command header.

TABLE I
GPSRQ PACKET HEADER FIELDS USED IN RECOVERY MODE FORWARDING

Field	Description
inRec	Mode: Greedy or Recovery
recPosition	IP address of the node where recovery mode started
loop	Returning Loop indicator
recIF	Interface used for forwarding in recovery mode

C. GPSRQ Protocol Implementation

To make the routing operation easier, encryption and authentication are performed on the data link TCP/IP layer [17], [73]. However, in previously deployed QKD networks the QKD header was transported unencrypted [74]. A denial-of-service attack is possible if an eavesdropper is able to block high importance packets such as routing packets or post-processing operation packets which would disable the link [75] or divert traffic to the node under an attacker's control [34]. Therefore, we propose using the QKD Command header, which is encrypted together with the packet payload (Fig. 6). Instead of adding the GPSRQ header to the packet, which would increase the amount of key used for encryption, values from the GPSRQ header which are listed in Table I are rearranged to the QKD header and QKD command header before transmission. Upon receiving the packet, the values are moved from the QKD header and QKD command header to the GPSRQ header.

In typical multimedia communication protocols, excessively delayed packets are not used for the reconstruction of transmitted data at the receiver [76]. We propose the extension of the QKD header by including *Timestamp* and *MaxDelay* fields to carry the packet's timestamp and maximum tolerated delay, respectively. The values of these fields are written at the ingress node, while the GPSRQ at the interior node checks the values before forwarding the packet, including the previously mentioned field "loop". If the value of the "loop" field is equal to 2, the packet was in a loop which was later avoided and marked in the internal cache in the previous nodes, so no action is taken against the packet. If the value of the "loop" field is equal to 0, it means that the packet was not previously in the loop. Then, GPSRQ checks whether the delay of the packet, which is calculated as the difference between the current timestamp and the Timestamp value, is greater than the MaxDelay. If this is true, GPSRQ sets the "loop" field to 1 and returns the packet to the previous-hop which will result in adding a new entry to the internal cache in previous nodes as mentioned earlier. The function of the Timestamp and MaxDelay fields is similar to the Time-to-Live (TTL) field in the IP header of conventional networks, but with the aim of minimizing the consumption of scarce key material in QKD network.

VII. Simulation Setup

To ensure simulations are independent of the topology or characteristics of any specific network, random graphs were constructed. Random network topologies were generated using the Waxman model [77], which is recommended for small and medium-size networks which include locality aspects such as QKD networks [2], [6], [49]. Additionally, the Waxman model corresponds to the requirement for the implementation of QKD networks without hierarchical multi-plane organization [78] since it spreads nodes randomly on a grid and adds links randomly, such that the probability P_e of interconnecting two nodes in a single plane is parameterized by the Euclidean distance separating them, as defined using Equation (12):

$$P_e(u,v) = \Theta \cdot exp^{-\frac{d(u,v)}{\Omega \cdot \Lambda}} \tag{12}$$

where $d(u,v)$ is the Euclidean distance between nodes u and v, Λ denotes the maximum possible distance between two nodes where $0 < \Omega, \Theta \leq 1$. The parameter choices are constrained to assure $P_e(u,v) \in [0,1]$. In total, 1290 simulations were performed with 30 randomly generated simulation seeds defining random network topologies and random values of the initial amount of key material in key storages. We evaluated GPSRQ ($\beta = 0.6$; $T_{average} = 5$) against OSPFv2 which was used in previously deployed QKD networks [17], [26], [39]–[41], and against DSDV which was used in our previous work [75]. The simulation was performed using the QKD Network Simulation Model (QKDNetSim) [79] to deploy GPSRQ and DSDV while NS-3-DCE v.1.9 was used to deploy the OSPFv2 routing protocol [80]. We used the BRITE topology generator to generate random topologies according to the Waxman model since it is supported under NS-3 and the source code is freely available [81]. NS-3-DCE and QKDNetSim were set to share the same seed file to generate random values which enabled us to use the same random topologies with the same configuration values in QKDNetSim and NS-3- DCE. Simulations included static networks with 10, 20, 30, 40, 50, 60 and 70 nodes which were randomly placed in a rectangular region and connected to QKD links with the following settings: minimal amount of key material 1 Mbyte; maximal amount 100 Mbyte; initial amount randomly generated in the [0.5, 25] MByte range; maximal bandwidth of the link set to 10 Mbps; charging key rate set to 100 kbps with the charging key period of seven seconds. The Waxman router model was used with the following parameter values: $m = 2$; $\Lambda = 100$; $\Omega = 0.4$; $\Theta = 0.4$; The first node was set as the source of traffic, while the last randomly placed node was set as the destination. The source node generated UDP traffic with a 1 Mbps rate and fixed packet size of 512 bytes which were encrypted using OTP and authenticated using VMAC with a 32-bit authentication tag [27]–[29]. The duration of the simulation was 150 seconds, while the capacity of waiting queues per device on the L2 and L4 layers (used for GPSRQ only) was set to 1000 packets. The parameters not given here are the default parameters of QKDNetSim and NS-3-DCE. In following sections, simulation results are presented using box-plot graphs.

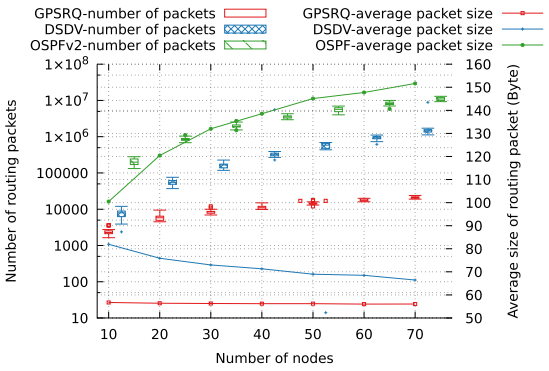

Fig. 7. The number and average sizes of routing packets.

A. Simulation Results and Evaluation

Although GPSRQ can achieve valuable results on non-planar graphs, one of the drawbacks of GPSRQ is the massive consumption of key material in non-planar networks since geographical routing cannot quickly determine the shortest path towards the destination which leads to unnecessary forwarding. It is, therefore, advisable to convert non-planar graphs into planar graphs in which geographic routing protocols can be effectively used. Instead of using the heuristic to exclude the intersecting edges [82], we modified the BRITE random topology generator to generate random planar Gabriel Graphs[7] (GG) which are defined as follows: "*An edge* (u,v) *exists between vertices* u *and* v *if no other vertex* w *is present within the circle whose diameter is the Euclidean distance d(u,v)*" [83], [84].

1) Routing Protocol Overhead: Fig. 7 shows the routing protocol overhead, measured by the number and average sizes

[7]The BRITE random generator with Gabriel Graph module is available at *https://bitbucket.org/mickeyze/brite-planar-graph*

of routing packets sent network-wide during the simulation for GPSRQ, DSDV and OSPFv2. OSPFv2 is a widely deployed link-state routing protocol which uses the periodic Link State Announcement (LSA) flooding mechanism to update link-state databases describing the network topology [85]. By default, OSPFv2 floods LSA update information every 30 minutes, and it exchanges Hello packets to establish and maintain a neighbor relationship every 10 seconds. If a node does not receive a Hello message from a neighbor within a fixed dead interval of time which is set to a default of 40 seconds for point-to-point networks, OSPFv2 modifies its topology database entries to indicate that the neighbor is unavailable. DSDV is a proactive routing protocol which periodically broadcast its routing table to its neighbors (every 15 seconds by default). Besides, DSDV uses triggered updates when the network topology suddenly changes. With the increase of the number of network nodes, the number of triggered packets is rising, which makes the average size of DSDV packets to decline [60], [86]. In contrast, GPSRQ relies on the knowledge of the geographical position and state of links to neighbors, which provides a high level of network scalability. As such, it periodically exchanges only M_{thr} packets defined in Section V-A every time new key material is stored in key storages; the value was set to seven seconds in our simulations. It is important to note that DSDV and OSPFv2 exchange their routing packets using UDP, while GPSRQ exchanges M_{thr} values using TCP, so Fig 7 shows all packets including TCP SYN, TCP ACK and TCP FIN. We performed several simulations where M_{thr} was set to 5, 15, 20 and 30 seconds besides the exchange of M_{thr} every time new key material is added to key storage. The data showed the same values, letting us conclude that a single exchange of a M_{thr} value in a QKD post-processing period is adequate for the smooth operation of GPSRQ. In our simulations, values from GPSRQ packets were moved into the QKD command header as described in Section VI-C, while DSDV and OSPFv2 packets were sent using the standard QKD header [17]. Fig 7 shows that GPSRQ consumes the lowest amount of key material for cryptographic operations in routing packets, while Fig. 9 shows that GPSRQ consumes the greatest amount of the available key material, but for securing application traffic as seen from the PDR value (Section VII-A2).

2) Packet Delivery Ratio: The Packet Delivery Ratio (PDR), calculated as the ratio of received and sent application packets, is used to assess the effectiveness of the routing protocol within the specified simulation environment. Fig. 8 shows that GPSRQ is able to successfully find available routes to the destination when compared to OSPFv2 and DSDV. Due to the significant value of the dead interval, OSPFv2 is not able to react quickly to the changes in network topology, which results in a reduced PDR value. DSDV exchanges routing packets more often; this returns a higher PDR value, although it is still significantly lower than GPSRQ. It is important to note that in some simulation scenarios PDR could not reach value "1" because it depends on the network conditions as defined by random values, such as the initial amount of key material in key storages. More precisely, because of the randomly assigned amount of the initial key material in key buffers,

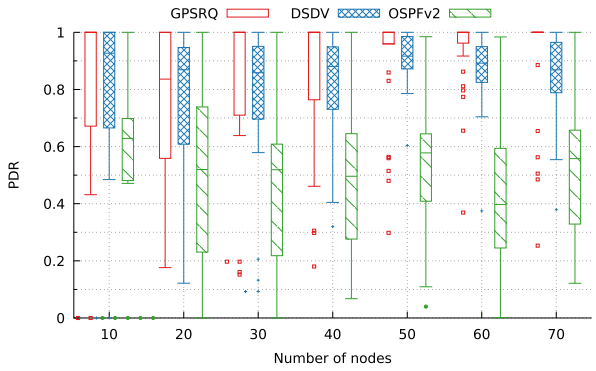

Fig. 8. Packet Delivery Ratio - the ratio of received and sent application packets shows that GPSRQ achieves significantly better results for networks with a larger number of nodes.

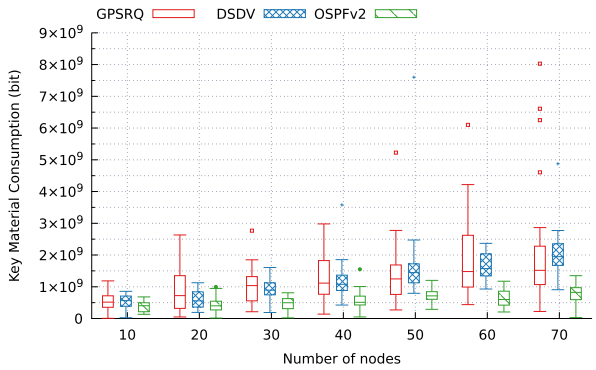

Fig. 9. Key material consumption - GPSRQ consumes more material because it successfully routes traffic to the destination.

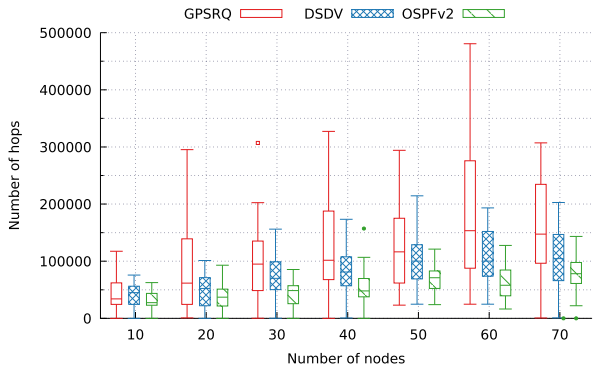

Fig. 10. Number of hops - GPSRQ finds paths to destinations that are not visible to other protocols.

several simulated scenarios were generated with little or no key available to establish the path to the destination regardless of which routing protocol is being used. However, using the same seed files ensures identical network conditions for all tested protocols in all simulated scenarios. By increasing the number of nodes, there is also a growing number of possible routes to the destination, which suits the GPSRQ protocol to achieve significantly better results than the other two protocols.

3) Path Length: Fig. 10 shows that GPSRQ recognizes paths to the destination which are not visible to DSDV and OSPFv2. More specifically, GPSRQ with the detection mechanism of the returning loop increases the number of

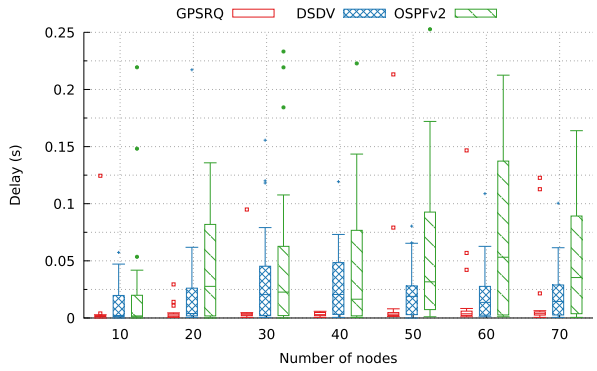

Fig. 11. End-to-end delay - GPSRQ caching techniques provide scalability and robustness resulting in the low and consistent delay.

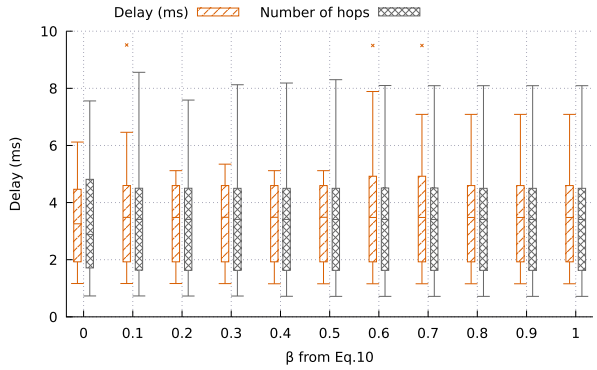

Fig. 12. Scaled average number of hops and scaled average delay for various values of β from Equation (10).

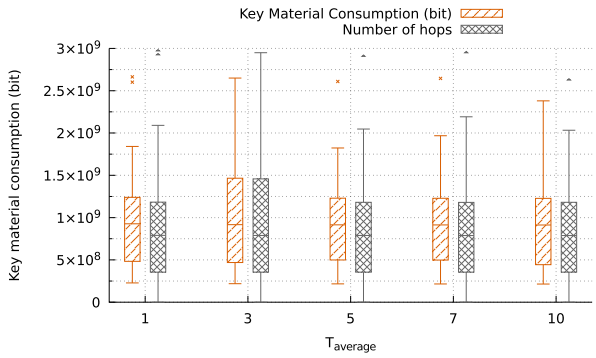

Fig. 13. Scaled average number of hops and scaled key material consumption for various values of $T_{average}$.

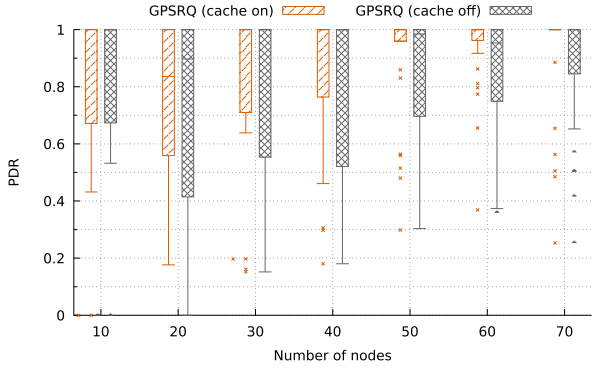

Fig. 14. The influence of the caching mechanism on PDR. $T_{average} = 5$.

forwarding instances. Still, due to scalable caching mechanism, GPSRQ is able to find the shortest feasible path to the destination.

4) End-to-End Delay: Our results show that the scalable caching in GPSRQ has a significant impact on the end-to-end delay. As shown in Fig. 11, values are significantly lower when compared to DSDV and OSPFv2 due to the active measurements of link states and detection of returning loops to exclude those links that do not lead to the destination. The end-to-end delay for GPSRQ does not notably increase with the number of nodes, leading to the conclusion that GPSRQ supports network scalability and robustness.

5) The Impact of GPSRQ β Parameter: GPSRQ uses Equation (10) to balance between the geographically shortest path and the path with the best performance. To show the impact of parameter β, we conducted simulations on 30 randomly generated network topologies with 30 randomly placed nodes, changing the values of β. Fig. 12 shows that parameter β, which is used in greedy forwarding has an impact on the number of hops and end-to-end delay. Considering that GPSRQ implements recovery mode as a side algorithm which continually seeks to forward packets to the destination using the right-hand rule, there are no significant differences in the PDR value for different values of β. However, β affects the number of hops to the destination in greedy forwarding, which results in an increased delay and overall consumption of key material. Intuitively, different classes of traffic should be served with varying values of β depending on the traffic priority.

6) Impact of Caching Mechanism: One of the key parameters of GPSRQ is the number of samples taken for calculating the average duration of key material establishment process in the long run, denoted as $T_{average}$ in Equation (8). $T_{average}$ defines the response time to variations of the public channel state, and it is used to determine the validity of entries in the GPSRQ cache memory T_{cache} as defined by Equation (11). $T_{average}$ affects the rate of emptying the internal node's cache, which has an impact in the number of hops to the destination. To demonstrate the impact of this parameter, we simulated 30 randomly generated network topologies with 30 randomly placed nodes, changing the values of $T_{average}$. Fig. 13 shows that the average number of hops as well as key consumption is related to the $T_{average}$. For small values, the GPSRQ caching mechanism does not provide significant gain since the cache memory is quickly cleared. For higher values, records in the cache memory are kept longer, resulting in a more stable network topology, which is reflected in the reduction of packet forwarding (number of hops).

To demonstrate the importance of the cache mechanism, we repeated experiments with the cache mechanism excluded. Fig. 14 shows the comparison of the PDR values for GPSRQ when cache mechanism is inactive and active. When the caching mechanism is turned off, the number of returning packets (packets with the value 2 of the loop field in the GPSRQ header) increases considerably. The key material is

TABLE II
QKD COMMAND HEADER - DESCRIPTION OF FIELDS

Field	Length	Short Description
Protocol	16 bits	Type of next header in the packet headers chain
Command	16 bits	Key management sub-protocol operation command (LOAD NEW KEY and other.)
RecIF	16 bit	GPSRQ recIf field
RecPosition	16 bit	GPSRQ recPosition field

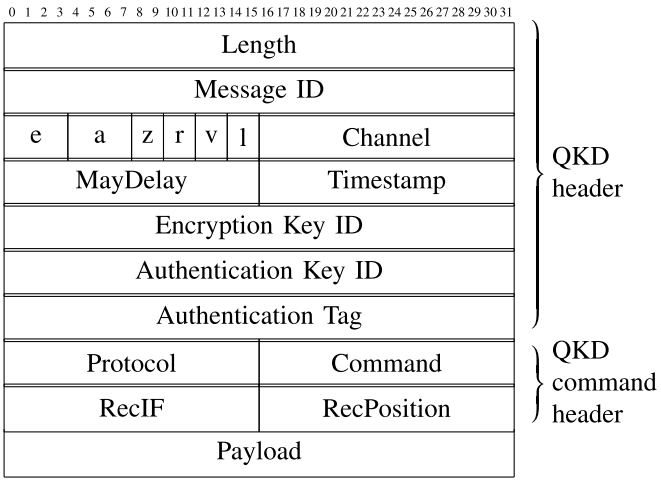

Fig. 15. Modified QKD header and QKD command header.

consumed on returning packets instead of data traffic which should be sent on the shortest path. Due to the larger traffic volume, the overall consumption of key material is substantially increased, which reduces the value of the PDR. Additionally, a noticeable increase in delay was noted due to longer routes.

VIII. CONCLUSION

This paper provides a novel QoS model and routing protocol for QKD networks. The FQKD QoS model involves traffic classification at the ingress node based on prioritizing traffic into appropriate queues. It also implements additional waiting queues at a higher network layer to adapt to the dynamic nature of QKD networks. It defines specific metrics for public and quantum channels, and overall QKD link metric and M_{thr} values to learn about the state of links which are more than one hop away. The GPSRQ routing protocol uses distributed geography and reactive routing to achieve high-level scalability. It is equipped with a caching mechanism and detection of returning loops, enabling forwarding while minimizing key material consumption. However, GPSRQ applications are limited to planar topologies only since geographic routing in networks with non-planar topologies are not able to quickly determine the shortest path, leading to unnecessary forwarding and increased consumption of scarce key material. Since QKD networks are limited to metropolitan scales [2], [6], [49], and previously deployed QKD networks were deployed on planar topologies [2]–[6], we do not consider this restriction as a critical deficiency. Our simulation results show that GPSRQ outperforms most popular solutions in previously deployed testbeds. It is reflected mainly by minimizing the average delay and increasing PDR, which are the key parameters for operative use of the real-time application [87].

The main contribution of this paper is providing novel metrics for determining the states of quantum and public channels as well as the overall state of QKD links, and providing a novel QoS model and routing protocol for QKD networks.

TABLE III
MODIFIED QKD HEADER - DESCRIPTION OF FIELDS

Field	Length	Short Description
Length	32 bits	Total packet length in bytes
Message ID	32 bits	Message ID
e	4 bits	Type of encryption cipher used
a	4 bits	Type of authentication algorithm used
z	2 bits	Type of compression algorithm used
v	2 bits	Version
r	2 bits	GPSRQ inRec indicator
l	2 bits	GPSRQ loop indicator
Channel	16 bits	QKD public channel ID
MaxDelay	16 bits	Maximum tolerated time delay
Timestamp	16 bits	Timestamp of the packet generation at the ingress node
Encryption Key ID	32 bits	ID of key used for Encryption
Authentication Key ID	32 bits	ID of key used for Authentication
Authentication-tag	32 bits	Authentication tag
Payload	-	Data payload

APPENDIX
MODIFIED QKD HEADER AND QKD COMMAND HEADER

Table II and Table III provide a short explanation of all fields in modified QKD headers and QKD command headers [17].

REFERENCES

[1] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proc. IEEE Int. Conf. Comput., Syst. Signal Process.*, vol. 175. New York, NY, USA, 1984, p. 8.
[2] M. Peev *et al.*, "The SECOQC quantum key distribution network in Vienna," *New J. Phys.*, vol. 11, no. 7, Jul. 2009, Art. no. 075001.
[3] C. Elliott and H. Yeh, "DARPA quantum network testbed," Raytheon BBN Technol. Cambridge, MA, USA, Tech. Rep., Jul. 2007.
[4] F. Xu *et al.*, "Field experiment on a robust hierarchical metropolitan quantum cryptography network," *Chin. Sci. Bull.*, vol. 54, no. 17, pp. 2991–2997, 2009.
[5] M. Sasaki *et al.*, "Field test of quantum key distribution in the Tokyo QKD network," *Opt. Express*, vol. 19, no. 11, pp. 10387–10409, Aug. 2011.
[6] S. Wang *et al.*, "Field and long-term demonstration of a wide area quantum key distribution network," *Opt. Express*, vol. 22, no. 18, pp. 21739–21756, Sep. 2014.
[7] A. Ciurana, V. Martin, J. Martinez-Mateo, B. Schrenk, M. Peev, and A. Poppe, "Entanglement distribution in optical networks," *IEEE J. Sel. Topics Quantum Electron.*, vol. 21, no. 3, pp. 37–48, May 2015.
[8] D. Collins, N. Gisin, and H. De Riedmatten, "Quantum relays for long distance quantum cryptography," *J. Modern Opt.*, vol. 52, no. 5, pp. 735–753, 2005.
[9] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, "Quantum repeaters based on entanglement purification," *Phys. Rev. A, Gen. Phys.*, vol. 59, no. 1, pp. 169–181, Jan. 1999.
[10] Z.-S. Yuan, Y.-A. Chen, B. Zhao, S. Chen, J. Schmiedmayer, and J.-W. Pan, "Experimental demonstration of a BDCZ quantum repeater node," *Nature*, vol. 454, no. 7208, pp. 1098–1101, Aug. 2008.
[11] L. Salvail, M. Peev, E. Diamanti, R. Alléaume, N. Lütkenhaus, and T. Länger, "Security of trusted repeater quantum key distribution networks," *J. Comput. Secur.*, vol. 18, no. 1, pp. 61–87, Jan. 2010.
[12] S. J. van Enk, J. I. Cirac, and P. Zoller, "Photonic channels for quantum communication," *Science*, vol. 279, no. 5348, pp. 205–208, Jan. 1998.

[13] R. Alléaume *et .al*, "Using quantum key distribution for cryptographic purposes: A survey," *Theor. Comput. Sci.*, vol. 560, pp. 62–81, Dec. 2014.

[14] A. A. V. A. Sergienko, *Quantum, Cryptography Communication*. Boca Raton, FL, USA: CRC Press, 2005.

[15] C. Elliott, "Building the Quantum Network," *New J. Phys.*, vol. 4, no. 1, p. 46, Jul. 2002.

[16] M. Marhoefer *et al.*, "Applicability of quantum cryptography for securing mobile communication networks," *Long-Term Dyn. Aspects Inf. Secur., Emerg. Trends Inf. Commun. Secur.*, 2007, pp. 97–111.

[17] C. Kollmitzer *et al.*, *Application Quantum Cryptography*, vol. 797. Cham, Switzerland: Springer, 2010,

[18] M. Dianati and R. Alleaume, "Architecture of the secoqc quantum key distribution network," in *Proc. 1st Int. Conf. Quantum, Nano, Micro Technol. (ICQNM)*, Jan. 2007, p. 13.

[19] R. Alleaume *et al.*, "Topological optimization of quantum key distribution networks," *New J. Phys.*, vol. 11, no. 7, Jul. 2009, Art. no. 075002.

[20] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," *Rev. Modern Phys.*, vol. 74, no. 1, pp. 145–195, Mar. 2002.

[21] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," *Rev. Mod. Phys.*, vol. 81, no. 3, pp. 1301–1350, Sep. 2009.

[22] A. Boaron *et al.*, "Secure quantum key distribution over 421 km of Optical Fiber," *Phys. Rev. Lett.*, vol. 121, no. 19, Nov. 2018, Art. no. 190502.

[23] A. R. Dixon, Z. L. Yuan, J. F. Dynes, A. W. Sharpe, and A. J. Shields, "Continuous operation of high bit rate quantum key distribution," *Appl. Phys. Lett.*, vol. 96, Apr. 2010, Art. no. 161102.

[24] B. Korzh *et al.*, "Provably secure and practical quantum key distribution over 307 km of optical fibre," *Nature Photon.*, vol. 9, no. 3, pp. 163–168, 2015.

[25] S. Wang *et al.*, "2 GHz clock quantum key distribution over 260 km of standard telecom fiber," *Opt. Lett.*, vol. 37, no. 6, pp. 1008–1010, 2012.

[26] M. Dianati, R. Alléaume, M. Gagnaire, and X. Shen, "Architecture and protocols of the future European Quantum key distribution network," *Secur. Commun. Netw.*, vol. 1, no. 1, pp. 57–74, Jan. 2008.

[27] A. Abidin and J.-Å. Larsson, "Security of authentication with a fixed key in quantum key distribution," 2011, *arXiv:1109.5168*. [Online]. Available: https://arxiv.org/abs/1109.5168

[28] C. Portmann, "Key recycling in authentication," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4383–4396, Jul. 2014.

[29] M. N. Wegman and J. L. Carter, "New hash functions and their use in authentication and set equality," *J. Comput. Syst. Sci.*, vol. 22, no. 3, pp. 265–279, Jun. 1981.

[30] M. Mehic, M. Niemiec, and M. Voznak, "Calculation of the key length for quantum key distribution," *Elektronika Ir Elektrotechnika*, vol. 21, no. 6, pp. 81–85, Dec. 2015.

[31] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris, "Resilient overlay networks," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 32, no. 1, p. 66, Jan. 2002.

[32] G. M. Lee and T. Choi, "Improving the interaction between overlay routing and traffic engineering," in *NETWORKING 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet* (Lecture Notes in Computer Science), vol. 4982, A. Das, H. K. Pung, F. B. S. Lee, and L. W. C. Wong, Eds. Berlin, Germany: Springer, 2008, doi: 10.1007/978-3-540-79549-0_46.

[33] Y. Liu, H. Zhang, W. Gong, and D. Towsley, "On the interaction between overlay routing and underlay routing," in *Proc. IEEE 24th Annu. Joint Conf. IEEE Comput. Commun. Societies*, vol. 4, Mar. 2005, pp. 2543–2553.

[34] S. Rass and S. König, "Turning Quantum Cryptography against itself: How to avoid indirect eavesdropping in quantum networks by passive and active adversaries," *Int. J. Adv. Syst. Meas.*, vol. 5, no. 1, pp. 22–33, 2012.

[35] Q. C. Le, P. Bellot, and A. Demaille, "Towards the world-wide quantum network," in *Proc. Int. Conf. Inf. Secur. Pract. Exper.* Berlin, Germany: Springer, 2008, pp. 218–232, doi: 10.1007/978-3-540-79104-1_16.

[36] C. L. Quoc, P. Bellot, and A. Demaille, "Stochastic routing in large grid-shaped quantum networks," in *Proc. IEEE Int. Conf. Res., Innov. Vis. Future (RIVF)*, Mar. 2007, pp. 166–174.

[37] S. Rass, "On game-theoretic network security provisioning," *J. Netw. Syst. Manage.*, vol. 21, no. 1, pp. 47–64, Mar. 2013.

[38] C. Elliott, A. Colvin, D. Pearson, O. Pikalo, J. Schlafer, and H. Yeh, "Current status of the DARPA quantum network," *Proc. SPIE, Quantum Inf. Comput. III, Int. Soc. Opt. Photon.*, vol. 5815, pp. 138–149, May 2005, doi: 10.1117/12.606489.

[39] Y. Tanizawa, R. Takahashi, and A. R. Dixon, "A routing method designed for a quantum key distribution network," in *Proc. 8th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2016, pp. 208–214.

[40] Y. Sun *et al.*, "Quality of service realization method applied to quantum key distribution network," CN 102 394 745 B, Dec. 2012. [Online]. Available: https://patents.google.com/patent/CN102394745B/en?q=%22Quality+of+service+realization+method+applied+to+quantum+key+distribution+network%22&oq=%22Quality+of+service+realization+method+applied+to+quantum+key+distribution+network%22

[41] C. Xianzhu, Y. Sun, and Y. Ji, "A QoS-supported scheme for quantum key distribution," in *Proc. Int. Conf. Adv. Intell. Awareness Internet (AIAI)*, 2011, pp. 220–224.

[42] M. Mehic, O. Maurhart, S. Rass, D. Komosny, F. Rezac, and M. Voznak, "Analysis of the public channel of quantum key distribution Link," *IEEE J. Quantum Electron.*, vol. 53, no. 5, Oct. 2017, Art. no. 9300408.

[43] P. Fazio, F. De Rango, and C. Sottile, "A predictive cross-layered interference management in a multichannel MAC with reactive routing in VANET," *IEEE Trans. Mobile Comput.*, vol. 15, no. 8, pp. 1850–1862, Aug. 2016.

[44] P. Fazio, M. Tropea, F. De Rango, and M. Voznak, "Pattern prediction and passive bandwidth management for hand-over optimization in QoS cellular networks with vehicular mobility," *IEEE Trans. Mobile Comput.*, vol. 15, no. 11, pp. 2809–2824, Nov. 2016.

[45] F. De Rango, P. Fazio, F. Scarcello, and F. Conte, "A new distributed application and network layer protocol for VoIP in mobile ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 10, pp. 2185–2198, Oct. 2014.

[46] F. De Rango, F. Guerriero, and P. Fazio, "Link-stability and energy aware routing protocol in distributed Wireless networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 4, pp. 713–726, Apr. 2012.

[47] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, "RSVP: A new resource ReSerVation Protocol," *IEEE Netw.*, vol. 7, no. 5, pp. 8–18, Sep. 1993.

[48] A. Ciurana *et al.*, "Quantum metropolitan optical network based on wavelength division multiplexing," *Opt. Express*, vol. 22, no. 2, pp. 1576–1593, 2014, doi: 10.1364/OE.22.001576.

[49] A. Ciurana *et al.*, "Quantum metropolitan optical network based on wavelength division multiplexing," *Opt. Express*, vol. 22, no. 2, pp. 1576–1593, 2014.

[50] T.-Y. Chen *et al.*, "Metropolitan all-pass and inter-city quantum communication network," *Opt. Express*, vol. 18, no. 26, pp. 27217–27225, 2010.

[51] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, "Experimental quantum cryptography," *J. Cryptol.*, vol. 5, no. 1, pp. 3–28, Jan. 1992.

[52] M. Mehic, P. Partila, J. Tovarek, and M. Voznak, "Calculation of key reduction for B92 QKD protocol," *Proc. SPIE, Quantum Inf. Comput. XIII, Int. Soc. Opt. Photon.*, vol. 9500, p. 95001J, May 2015, doi: 10.1117/12.2177149.

[53] M. Mehic, A. Maric, and M. Voznak, "QSIP: A quantum key distribution signaling protocol," in *Proc. Int. Conf. Multimedia Commun., Services Secur.* Cham, Switzerland: Springer, 2017, pp. 136–147, doi: 10.1007/978-3-319-69911-0_1.

[54] O. Maurhart, T. Lorunser, T. Langer, C. Pacher, M. Peev, and A. Poppe, "Node modules and protocols for the quantum-back-bone of a quantum-key-distribution network," in *Proc. 35th Eur. Conf. Opt. Commun.*, Sep. 2009, pp. 3–4.

[55] Y. Zhu, C. Dovrolis, and M. Ammar, "Dynamic overlay routing based on available bandwidth estimation: A simulation study," *Comput. Netw.*, vol. 50, no. 6, pp. 742–762, Apr. 2006.

[56] D. G. Andersen, A. C. Snoeren, and H. Balakrishnan, "Best-path vs. Multi-path overlay routing," in *Proc. ACM SIGCOMM Conf. Internet Meas. (IMC)*, Oct. 2003, pp. 91–100.

[57] J. Cederlof and J.-Å. Larsson, "Security aspects of the authentication used in quantum cryptography," *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1735–1741, Apr. 2008.

[58] D. Dodson *et al.*, "Updating quantum cryptography report ver. 1," 2009, *arXiv:0905.4325*. [Online]. Available: https://arxiv.org/abs/0905.4325

[59] A. Sergienko, S. Pascazio, and P. Villoresi, *Quantum Communication and Quantum Networking*. Berlin, Germany: Springer, 2010, doi: 10.1007/978-3-642-11731-2.

[60] M. Mehic *et al.*, "On using multiple routing metrics with destination sequenced distance vector protocol for MultiHop wireless ad hoc networks," *Proc. SPIE, Model. Simul. Defense Syst. Appl. XI, Int. Soc. Opt. Photon.*, vol. 9848, p. 98480F, May 2016, doi: 10.1117/12.2223671.

[61] J. Li, J. Jannotti, D. S. J. De Couto, D. R. Karger, and R. Morris, "A scalable location service for geographic ad hoc routing," in *Proc. 6th Annu. Int. Conf. Mobile Comput. Netw.*, Aug. 2000, pp. 120–130.
[62] S. Wijesekera, "Quantum cryptography for secure communication in IEEE 802.11 wireless networks," Ph.D. dissertation, Dept. Inf. Sci. Eng., Univ. Canberra, Canberra, ACT, Australia, 2011.
[63] K. H. Sheikh, S. S. Hyder, and M. M. Khan, "An overview of quantum cryptography for wireless networking infrastructure," in *Proc. Int. Symp. Collaborative Technol. Syst., (CTS)*, May 2006, pp. 379–385.
[64] T. Schmitt-Manderbach *et al.*, "Experimental demonstration of free-space decoy-state quantum key distribution over 144 km," *Phys. Rev. Lett.*, vol. 98, Jun. 2007, Art. no. 010504.
[65] G. Vallone *et al.*, "Experimental satellite quantum communications," *Phys. Rev. Lett.*, vol. 115, no. 4, Jul. 2015, Art. no. 040502.
[66] J. Yin *et al.*, "Satellite-based entanglement distribution over 1200 kilometers," *Science*, vol. 356, no. 6343, pp. 1140–1144, 2017.
[67] M. Dusek *et al.*, "Quantum Cryptography," in *Progress in Optics*, vol. 49. Amsterdam, The Netherlands: Elsevier, Jan. 2006, pp. 381–454.
[68] S. K. Liao *et al.*, "Space-to-ground quantum key distribution using a small-sized payload on Tiangong-2 space lab," *Chin. Phys. Lett.*, vol. 34, no. 9, Aug. 2017, Art. no. 090302.
[69] C. J. Pugh *et al.*, "Airborne demonstration of a quantum key distribution receiver payload," *Quantum Sci. Technol.*, vol. 2, no. 2, Jun. 2017, Art. no. 024009.
[70] K. Shimizu *et al.*, "Performance of long-distance quantum key distribution over 90-km optical links installed in a field environment of Tokyo metropolitan area," *J. Lightw. Technol.*, vol. 32, no. 1, pp. 141–151, Jan. 1, 2014.
[71] S. Basagni *et al.*, "A distance routing effect algorithm for mobility (DREAM)," in *Proc. 4th Annu. ACMIEEE Int. Conf. Mobile Comput. Netw.*, Oct. 1998, pp. 76–84.
[72] B. Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks," in *Proc. 6th Annu. Int. Conf. Mobile Comput. Netw.*, Aug. 2000, pp. 243–254.
[73] M. Mehic, D. Komosny, O. Mauhart, M. Voznak, and J. Rozhon, "Impact of packet size variation in overlay quantum key distribution network," in *Proc. 11th Int. Symp. Telecommun. (BIHTEL)*. Sarajevo, Bosnia-Herzegovina, Oct. 2016, pp. 1–6.
[74] O. Maurhart *et al.*, "New release of an open source QKD software: Design and implementation of new algorithms, modularization and integration with IPSec," in *Proc. Qcrypt*, Aug. 2013, p. 1.
[75] M. Mehic, P. Fazio, M. Voznák, and E. Chromý, "Toward designing a quantum key distrubution network simulation model," *Adv. Electr. Electron. Eng.*, vol. 14, no. 4, pp. 413–420, 2016.
[76] L. Sun, I.-H. Mkwawa, E. Jammeh, and E. Ifeachor, *Guide to Voice and Video Over IP: For Fixed and Mobile Networks*. London, U.K.: Springer, 2013.
[77] B. M. Waxman, "Routing of multipoint connections," *IEEE J. Sel. Areas Commun.*, vol. 6, no. 9, pp. 1617–1622, Dec. 1988.
[78] W. Klaus *et al.*, *Modeling and Tools for Network Simulation*. Berlin, Germany: Springer, 2010.
[79] M. Mehic, O. Maurhart, S. Rass, and M. Voznak, "Implementation of quantum key distribution network simulation module in the network simulator NS-3," *Quantum Inf. Process.*, vol. 16, no. 10, p. 253, Oct. 2017.
[80] G. F. Riley and T. R. Henderson, *The ns-3 Network Simulator*. Berlin, Germany: Springer, 2010, pp. 15–34.
[81] A. Medina, A. Lakhina, I. Matta, and J. Byers, "BRITE: An approach to universal topology generation," in *Proc. MASCOTS 9th Int. Symp. Modeling, Anal. Simulation Comput. Telecommun. Syst.*, Aug. 2001, pp. 346–353.
[82] B. N. Karp, "Geographic routing for wireless networks," Ph.D. dissertation, Dept. Eng. Appl. Sci., Harvard Univ., Cambridge, MA, USA, 2000. [Online]. Available: https://www.comp.nus.edu.sg/~bleong/geographic/related/karp00geographic.pdf
[83] K. M. Chandy and J. Misra, "Distributed computation on graphs: Shortest path algorithms," *Commun. ACM*, vol. 25, no. 11, pp. 833–837, Nov. 1982.
[84] K. R. Gabriel and R. R. Sokal, "A new statistical approach to geographic variation analysis," *Syst. Biol.*, vol. 18, no. 3, pp. 259–278, 1969.
[85] J. T. Moy, *OSPF Version 2 Internet Requests for Comment*, document RFC 1247, 1991, pp. 1–124. [Online]. Available: https://tools.ietf.org/html/rfc1247
[86] E. C. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 24, no. 4, pp. 234–244, 1994.
[87] M. Voznak, A. Kovac, and M. Halas, "Effective packet loss estimation on VoIP jitter buffer," in *Proc. Int. Conf. Res. Netw.*, vol. 7291. Berlin, Germany: Springer, 2012, pp. 157–162, doi: 10.1007/978-3-642-30039-4_21.