Investigating the Effects of Packet Dropping and Packet Pollution Attacks in Network Coding Networks

H.L.H.C. Terblanche, M.J. Grobler and H. Marais
School of Electrical, Electronic and Computer Engineering
North-West University, Potchefstroom Campus
Tel: +27 18 299 4296, Fax: +27 18 299 1977
Email: {20569807, leenta.grobler, henri.marais}@nwu.ac.za

***Abstract*—Network Coding (NC) has introduced a new way to increase the efficiency of networks, especially Wireless Mesh Networks (WMNs). There are many security threats when using NC in WMNs. We investigate the effects packet pollution and packet dropping attacks in a NC environment. Packet pollution decreases the throughput of the network dramatically, while packet dropping does not have such a big effect. By adding security to NC the effects of these attacks can be decreased.**

***Index Terms*—MATLAB®, Network Coding, Packet Dropping, Packet Pollution, Security**

I. Introduction

Wireless Mesh Networks (WMNs) are multi-hop networks, with a decentralised nature, that can dynamically form into mesh topologies. They offer many advantages, such as low installation costs, easy maintenance, network robustness and reliable service coverage. As more efficient protocols are developed, these type of networks are becoming more popular. One way to improve the efficiency of these protocols is something called Network Coding (NC).

NC is a method used to increase the efficiency of networks by encoding and decoding data on packet level by means of a logical XOR operation [1]. NC works well in WMNs because it can exploit the broadcast and opportunistic listening properties of this type of network.

With WMN's increase in popularity, the issue of security became prominent. Because NC relies on the combination of valid packets to generate forwarded messages, the network becomes vulnerable to packet pollution attacks. There are numerous ways in which these attacks are addressed, [2]–[4]. Most of these schemes use homomorphic hash signatures which incur a lot of overhead [5] and diminishes the advantage gained by NC.

In 2009 Dong et al. [5] proposed a security scheme for NC in WMNs that claimed not to have as much overhead as previous schemes to address packet pollution. This scheme is based on time asymmetry and checksums. Another security scheme for peer-to-peer networks proposed in [6] is also based on similar principles.

In this paper we look at the effects of packet pollution and packet dropping in networks using NC.

In section II background will be given on WMNs, NC and security in NC, along with a discussion on the DART security scheme. In section III the simulator model and experimental setup will be described. After that, in sections IV and V the results will be discussed and a conclusion will be drawn.

II. Background

A. Wireless Mesh Networks

WMNs are wireless multi-hop networks that consist of wireless clients, routers and gateways. They support ad-hoc networking that has self-forming, self-healing and self-organization properties.

There are three different architectures: backbone or infrastructure mesh networks, client mesh networks and hybrid mesh networks [7].

- **Backbone or Infrastructure WMNs** are the most common. Mesh routers form a backbone for mesh clients to connect to, and the routers can connect to the internet if it doubles as a gateway, as shown in Fig. 1.
- **Client WMNs** consist only of wireless clients that communicate with each other.
- **Hybrid WMNs** are a combination of client and backbone mesh networks.

The advantages of WMNs are [7]:

- Network robustness;
- Reliable service coverage;
- Easy maintenance;
- Low installation costs.

Despite these advantages, the decentralised nature and the openness of the medium, creates a security risk because there is no authentication and the network is vulnerable to eavesdropping. NC can address these problems to a certain degree, because the packets are encoded, and an eavesdropper must first get enough packets before it can get any real information.

B. Network Coding

As stated in the introduction, NC can improve the efficiency of a network. This is accomplished by forming linear

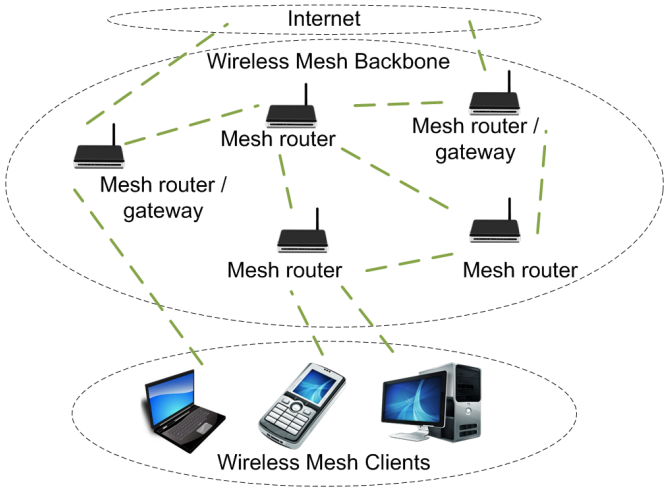

Fig. 1: Backbone Wireless Mesh Network

combinations of the received packets and forwarding only the combinations. A coding vector is attached to the new packet specifying which packets were combined to generate the encoded packet. These combined packets can easily be decoded at the receiver node. This technique was first proposed in 2000 by Ahlswede et al. in [8].

Random Linear Network Coding (RLNC) was first introduced by Ho et al. [9], where the elements in the coding vector are randomly chosen from a finite field, and the packets are combined accordingly. The most common fields that are used are the G_{2^8} and $G_{2^{16}}$ fields. It was proven by Ho et al. [9] that if the field is sufficiently large the resulting encoded packet will be linearly independent from the other native and encoded packets. A native packet is a packet that has not been encoded, while an innovative packet is a packet that is linearly independent from the other received native and encoded packets. In RLNC, the intermediate/forwarder nodes can decide which packets to combine, using the random coding vectors, before forwarding them. When enough innovative packets have arrived at the receiver it is easy to decode them using Gaussian elimination. With RLNC, there is no need for a centralised control mechanism and encoded packets do not have to arrive in sequence at the receiver.

The advantages of NC include robustness, maximizing the throughput, increasing the efficiency and minimizing the delay of the network [1], [10].

RLNC works well in WMNs as shown by Ho et al. in [11]. Although using NC with WMNs has many benefits it also introduces vulnerabilities to the network that has to be addressed.

C. Security in Network Coding

Security in NC was first addressed by [12] in 2002. There are three approaches to NC security [13]:

- Computational - when it is computationally infeasible to break the system;
- Physical - when using physical properties to prevent or detect attacks; and
- Information Theoretic - determines the maximum transmission rate necessary to make it impossible to break the system.

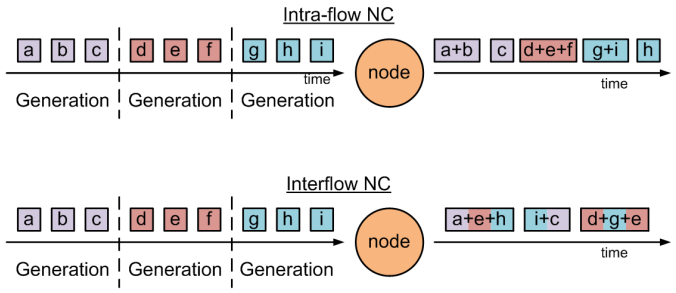

Fig. 2: Intra-flow and Inter-flow Network Coding

To ensure that the decoding process is optimised, the information that has to be sent, is divided into chunks of n packets, called generations. There are two general approaches for NC in WMNs - *intra-flow NC* and *inter-flow NC*. Intra-flow NC combines packets within individual generations and inter-flow NC combines packets across different generations as shown in Fig. 2.

In [14], Dong analysed the threats for both general approaches to NC. They divided each approach into different components. The components for intra-flow NC were *forwarding node selection and rate assignment*, *data packet forwarding* and *acknowledgement delivery*.

The threats identified were:

For forwarding node selection and rate assignment:

- Link quality falsification or modification;
- Wormholes.

For data packet forwarding:

- Packet pollution;
- Packet dropping.

For acknowledgement delivery:

- ACK injection or modification;
- ACK dropping;
- ACK delay.

In this paper we focus on the *data packet forwarding* component of intra-flow NC. For this study the other identified threats were not taken into account. Thus, we focus on packet pollution and packet dropping. Of the two, packet pollution has the highest impact on the throughput of a network using NC.

Packet Pollution: Packet pollution occurs when a malicious node injects corrupt packets into the network. Packet pollution can also occur accidentally when the packets get corrupt because of channel errors. Packet pollution attacks can cause a significant decrease in the throughput of the network.

Because packets are combined, the pollution can spread quickly down the network.

The packet pollution security threat has been addressed by other security schemes [2]–[4]. These schemes all incur high overhead. The DART scheme proposed by [5] addresses

packet pollution but does not incur as much overhead as other schemes.

Packet Dropping: The packets can be dropped by a malicious node or the packets can be lost because of channel errors. Packet dropping attacks where the malicious node drops all the packets received, are known as black hole attacks. With grey hole attacks packets are dropped at random intervals. Grey hole attacks are usually more difficult to detect than black hole attacks. Packet dropping attacks are not as severe as packet pollution but still has a negative effect on the throughput of the network.

D. DART Security Scheme

When implementing the DART security scheme the source divides the data into generations, and encoded packets are generated from the active generation and sent into the network. The source also generates checksum packets, at constant time intervals, for the active generation and broadcasts it to all the forwarder nodes.

The encoded packets that arrive at the forwarder nodes are stored in an unverified queue. These packets are then verified upon reception of a checksum packet and only packets that arrived at the node before the checksum was created are verified. Packets that pass verification are then put into the verified queue and packets that failed are discarded.

Packets that arrive at the receiver node also go through the same process as the forwarder nodes but the verified packets are stored in a decoding matrix. If the receiver node has enough innovative packets of the current generation, it decodes the generation and sends an acknowledgement packet to the sender to begin with the next generation.

If a malicious node injects a polluted packet into the network, it will not propagate further than one hop, because of the verification of packets at each node. If an attacker can produce a polluted packet that meets the current checksum's requirements, the packet still will not be verified because the node only verifies packets that were received before the checksum was created. The polluted packet will thus be verified by the next different random checksum and be discarded.

This security scheme was implemented in the Glomosim environment [15]. This environment is outdated, therefore it was decided to do a custom implementation in MATLAB®.

III. Methodology

A. Network Model

Consider a wireless network with one source node, one receiver node and *(0 - n)* forwarder nodes. Each node can send and receive packets. The network uses RLNC and only decodes packets at the receiver. It is a unicast network with the source connected to the receiver via the forwarder nodes as seen in Fig. 3.

B. Simulator Design/Model

1) Basic Network: Consider the simple network in Fig. 3. If the source wants to transmit the image in Fig. 4 to the receiver node, it has to perform the following steps. The source divides the file into generations. It creates an encoded packet from the current generation and sends it to the forwarder node. When the forwarder node receives the encoded packet, it stores it in a queue until it is ready to send. The forwarder node sends the packet to the next forwarder node where it also gets stored in a queue. The packet is then sent to the receiver. Upon reception the packet gets tested for linear independence and is stored in a decoding matrix. When the receiver node has received enough linearly independent packets of the current active generation, it decodes them using Gaussian elimination. After a generation has been decoded and verified the receiver sends an acknowledgement to the source node. Upon the reception of the acknowledgement packet, the source node begins sending packets from the next generation. This continues until there are no more generations left at the source node. If there were no faults the image should not be corrupt after decoding.

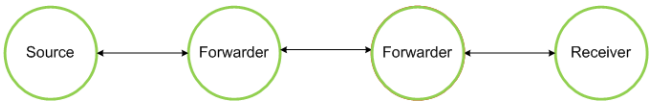

Fig. 3: Example of the network simulated

Fig. 4: The image at the source

2) Network with security enhancements: In the case of added security the source node generates a checksum packet for the current generation at fixed time intervals. The checksum packet is broadcast to all the first hop forwarder nodes. The forwarder and receiver nodes then use the checksum to verify packets. If they pass verification they are forwarded or stored in the decoding matrix, otherwise they are discarded.

When each generation is decoded it is verified by end-to-end authentication. If a decoded generation is corrupt no acknowledgement packet is sent and the receiver restarts packet collection.

C. Metrics

Throughput - The throughput of each network was measured to see what the effect of each attack was. The equation used is :

$$\text{Throughput} = \frac{\text{Total relevant encoded packets received}}{\text{Time to send file}} \quad (1)$$

Packet Loss - Lost packets are defined as encoded packets that are dropped or polluted by malicious nodes. Packet loss is determined by measuring the total relevant encoded packets received in a malicious environment, to the amount of relevant encoded packets received without any malicious node intervention.

Fig. 5: The image at the receiver after decoding

Fig. 6: Network with a malicious node present

D. Experimental Setup

The simulation was done in MATLAB® and time was measured in artificial ticks. The field used for NC was G_{2^8}, the generation size was 32 packets and the packet size was 1500 Bytes. The queue size for all the forwarder and receiver nodes was 10. For the security scheme the default setup in Dong [5] was used, with a checksum packet sent every 6 packets.

The image in Fig. 4 was sent through the network in Fig. 3 to establish a baseline of the throughput without any malicious nodes. The size of the image was 1.17 MB and was divided into 26 generations.

For an attacker scenario, the image was sent through the network in Fig. 6 with a malicious node present. A malicious node is a forwarder node that was modified to corrupt or drop some of the packets it forwards.

IV. Results

The effects of packet pollution and packet dropping were investigated for a simple network topology. The following graphs show the throughput and packet loss in networks with NC alone and then with NC with added security.

The image in Fig. 4 was sent through the network with only one polluted packet. When the data was decoded at the receiver the image was corrupt as seen in Fig. 5. From the fact that the picture was corrupt, when it arrived at the receiver, it was derived that the effective throughput was 0.

1) Baselines: Fig. 7 shows the throughput for the network in Fig. 3. Implementation of the security scheme in the same network causes a reduction in throughput as seen in Fig. 7. Upon further analysis of the graph it was seen that the reduction is about 30%.

2) No Security: Fig. 8 shows the throughput in cases of packet pollution with end-to-end verification. The graph shows that if only a single generation is polluted the throughput falls from the 89% of the baseline to 86%. The throughput decreases further as more generations are polluted. If all the generations have been polluted, by one packet per generation, then the throughput falls from the 89% baseline to 47%. It is assumed that a generation is only polluted once, meaning when it is sent the second time no pollution occurs.

In Fig. 8 the throughput for different cases of packet dropping is shown.

Case 1: 1 packet (3%) of a generation is dropped.

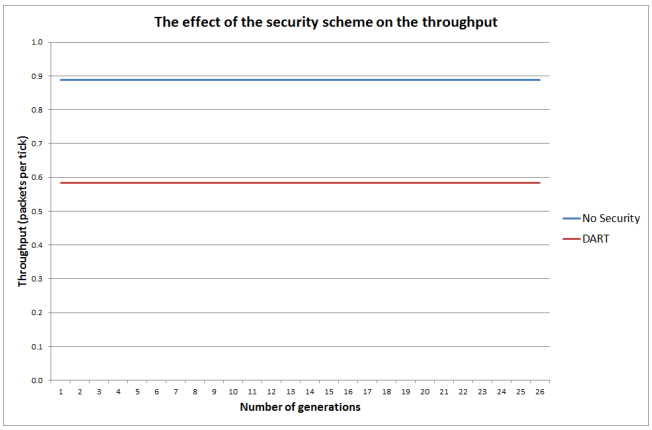

Fig. 7: Throughput of the network in Fig. 3

Case 2: 3 packets (10%) of a generation are dropped.

Case 3: 8 packets (25%) of a generation are dropped.

In case 1 the throughput decreases only slightly. If a packet were dropped in every generation, the throughput would decrease from 89% to 86%. In case 2, the throughput at 26 generations dropped from 89% to 82%. In case 3 the throughput decreased from 89% to 74% at 26 generations.

As can be seen from Fig. 8 and Fig. 10, packet pollution has a much larger effect on the throughput, of the network, than packet dropping.

3) DART Security: For attacks on the network with the DART security it was seen that the throughput of the network decreases only slightly as seen in Fig. 9. The same cases as above were applied for packet dropping. In case 1 the throughput decreases only slightly. If a packet was dropped in every generation the throughput would decrease from 59% to 57%. In case 2 the throughput at 26 generations dropped from 59% to 55%. In case 3 the throughput decreased from 59% to 51% at 26 generations. The packet dropping and packet pollution attacks did not have such a big effect on the packet loss, because of the nature of the security scheme. This can be seen in Fig. 11. The packet loss in the case for packet dropping and packet pollution, of 1 packet per generation, is the same. This is because the polluted packets are dropped when detected.

V. Conclusion

In this paper, we studied the effects of packet pollution and packet dropping, on a simple network using NC. We saw that if there is only one pollution attacker present in the network, the throughput decreases dramatically if each generation is polluted. In the case of a packet dropping attacker, the effect is not as great as with packet pollution. Packet pollution is the biggest threat to networks using NC. This can be seen in Fig. 8 where the packet pollution attack has a much greater effect on the throughput than the packet dropping attack, where more packets are dropped.

By implementing the DART security scheme the throughput was reduced, but no data corruption occurred due to packet pollution. The throughput, in the case where every generation

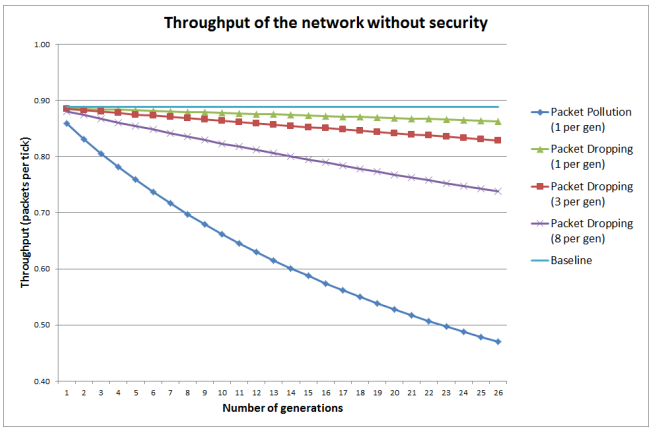

Fig. 8: Throughput of the network in Fig. 6 without security

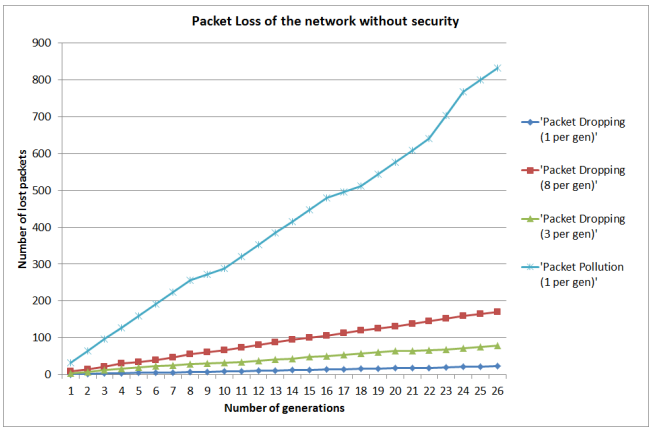

Fig. 10: Packet Loss of the network in Fig. 6 without security

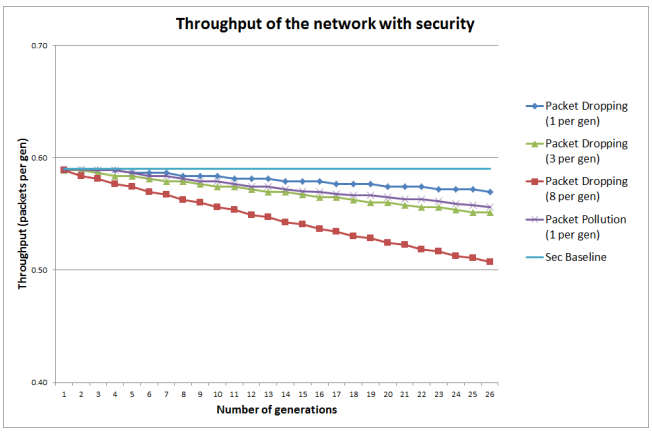

Fig. 9: Throughput of the network in Fig. 6 with DART security scheme

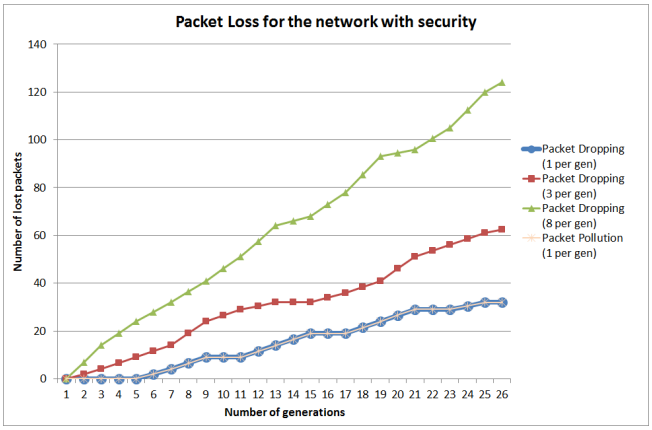

Fig. 11: Packet Loss of the network in Fig. 6 with DART security scheme

is polluted as seen in Fig. 8, is also lower than the effective throughput of the security baseline as seen in Fig. 7. Packet loss is also not as significant in the case of security as can be seen in Fig. 11.

VI. Future Work

Future work will include expanding the network and investigating how the security scheme, that already addresses packet pollution, can be improved to address packet dropping. This includes determining how the checksum packet can be used along with stored packets at forwarder nodes to recreate dropped packets in the network.

VII. Acknowledgements

This work was completed with funding from the Telkom Centre of Excellence at the NWU, Potchefstroom Campus.

References

[1] C. Fragouli, J.-Y. Le Boudec, and J. Widmer, "Network Coding: An Instant Primer," *ACM Sigcomm Computer Communication Review*, vol. 36, no. 1, pp. 63–68, 2006.

[2] T. Ho, B. Leong, R. Koetter, M. Medard, M. Effros, and D. Karger, "Byzantine modification detection in multicast networks using randomized network coding," in *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on DOI - 10.1109/ISIT.2004.1365180*, 2004, p. 144.

[3] A. Le and A. Markopoulou, "Locating byzantine attackers in intra-session network coding using spacemac," in *Network Coding (NetCod), 2010 IEEE International Symposium on DOI - 10.1109/NETCOD.2010.5487673*, 2010, pp. 1–6.

[4] S.-Y. R. Li, Q. T. Sun, and Z. Shao, "Linear network coding: Theory and algorithms," *Proceedings of the IEEE DOI - 10.1109/JPROC.2010.2093851*, vol. 99, no. 3, pp. 372–387, 2011.

[5] J. Dong and R. Curtmola, "Practical defenses against pollution attacks in intra-flow network coding for wireless mesh networks," in *Proceedings of the second ACM conference on Wireless network security*, ser. WiSec '09. New York, NY, USA: ACM, 2009, pp. 111–122.

[6] A. Le and A. Markopoulou, "Tesla-based defense against pollution attacks in p2p systems with network coding," in *Network Coding (NetCod), 2011 International Symposium on DOI - 10.1109/ISNETCOD.2011.5979096*, 2011, pp. 1–7.

[7] I. Akyildiz and X. Wang, "A survey on wireless mesh networks," *Communications Magazine, IEEE DOI - 10.1109/MCOM.2005.1509968*, vol. 43, no. 9, pp. S23–S30, 2005.

[8] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *Information Theory, IEEE Transactions on DOI - 10.1109/18.850663*, vol. 46, no. 4, pp. 1204–1216, 2000.

[9] T. Ho, R. Koetter, M. Medard, D. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," in *Information Theory, 2003. Proceedings. IEEE International Symposium on*, 2003, pp. 442–.

[10] P. Chou and Y. Wu, "Network coding for the internet and

wireless networks," *Signal Processing Magazine, IEEE DOI - 10.1109/MSP.2007.904818*, vol. 24, no. 5, pp. 77–85, 2007.
[11] T. Ho, B. Leong, M. Medard, R. Koetter, Y.-H. Chang, and M. Effros, "On the utility of network coding in dynamic environments," in *Wireless Ad-Hoc Networks, 2004 International Workshop on*, 2004, pp. 196–200.
[12] N. Cai and R. Yeung, "Secure network coding," in *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on DOI - 10.1109/ISIT.2002.1023595*, 2002, p. 323.
[13] N. Cai and T. Chan, "Theory of secure network coding," *Proceedings of the IEEE DOI - 10.1109/JPROC.2010.2094592*, vol. PP, no. 99, pp. 1–17, 2011.
[14] J. Dong, R. Curtmola, and C. Nita-Rotaru, "Secure network coding for wireless mesh networks: Threats, challenges, and directions," *Computer Communications*, vol. 32, no. 17, pp. 1790–1801, Nov. 2009.
[15] Glomosim. [Online]. Available: http://pcl.cs.ucla.edu/projects/glomosim/

H.L.H.C. Terblanche is a Telkom CoE student studying towards a Masters degree in Computer and Electronic Engineering at the NWU. She received her B.Eng in Computer and Electronic Engineering in 2010 from the NWU. Her current research interests are wireless mesh networks, network coding and security.