ISSN:2249-5789

Attacks on Phoenix: A Virtual Coordinate System

Manpreet Kaur
manpreet.7467@gmail.com
Lovely Professional University
Jalandhar (Punjab)

***Abstract*—** Many large-scale internet application take use of network coordinate systems to represent the network distance relationships among internet hosts. NC systems offer an efficient and scalable mechanism to predict the distance between internet hosts by mapping them into an appropriate geometrical space. This network distances are calculated in terms of round trip time (RTT). Most of the proposed systems were relays on Euclidean based model, which contains the problem of TIV. To overcome this problem, matrix factorization was introduced which has better prediction accuracy. However, the accuracy of such systems is often based on good cooperation of nodes. It is assumed that the nodes are altruistic and always report with correct information about coordinates.

But when a malicious node provides incorrect information the performance may degrade. In this paper, we reveal the susceptibility of virtual coordinate systems against the insider attacks. We demonstrate theses attacks on a well known coordinate system, Phoenix using simulation based on King's and PlanetLab. We show the impact of different malicious activities on the phoenix coordinate system.

***Keywords*—** Network Coordinates, Phoenix system, RTT, Triangular inequality violations, Anomaly Detection, Security, Reliability, Virtual coordinate system.

1. INTRODUCTION

Most of the internet application such as network monitoring [1], locality-aware server selection [2], application layer multicast [2], distributed query optimization [3], BitTorrent based file sharing [4], network modelling [5], and many more rely on the concept of network proximity, in terms of round trip times (RTTs). It is based on the pair-wise distance measurements among internet hosts, for appropriate neighbour selection. Tracking these measurements in this changing environment needs frequent measurements which consumes high rate of bandwidth.

Network coordinate systems (e.g. PIC [6], Vivaldi [7], Pharos [8], Phoenix [9], IDES [10], etc) are used to characterize the network distance (i.e. round trip time) among the internet hosts. For this, they assign practical coordinate to each and every host in a network by mapping them into an appropriate geometric space. Then it allows the estimations of network latencies between the every host in the network rather than having the direct measurements of end-to-end latencies.

This approach suits well in case of peer-to-peer applications, where nodes compute network distance by piggybacking the coordinate information. These systems provide the desirable properties such as robustness, stability, scalability along with accuracy and less measurement overheads.

However, these systems are exceptionally eminent to attacks. As they work on assumption that the nodes contributes in system cooperate with each other and always supply correct information. This manner of coordinate systems makes them open to nasty activities. Even an insider can also perform attacks by acting as a truthful neighbour.

In this paper, we learn such type of malicious activities on phoenix system. It is a decentralized system that does not require any fixed set of landmark nodes. We study the impact of coordination inflation, deflation and oscillation attacks on phoenix. The result is evaluated on the bases of relative error. In [18], these attacks are referred against VCS-routing. We will mitigate these attacks against phoenix system and check the performance of the system.

The rest of the paper is organized as follows. Section 2, provides a review of related work. In section 3, we present the various coordinates systems attacks. In section 4, we discuss the research methodology, followed by results evaluation in section 5. Section 6, is the conclusion.

2. RELATED WORK

In this section, we give a brief overview about the types of network coordinate systems, a description about the operational models of coordinate systems and describe the working of representative virtual coordinate system, Phoenix.

2.1 Classes of Network Coordinate Systems

The network coordinate systems are generally divided into two main classes: (a) the centralized systems and (b) decentralized systems.

(a) Centralized NC Systems

In centralized NCS, a fixed set of reference nodes are used to calculate the position of all the other nodes in system. These nodes are called the landmarks, which act as reference points for the calculating the coordinates of all other nodes. First of all, the coordinate position of the landmark nodes is calculated by reducing the error among the exact measured distance and the estimated distance of the landmark nodes, these coordinates are called the global coordinates of the system. After that, the ordinary nodes compute their coordinates by referring these Landmark nodes [12].

GNP [11], VL [13], ICS [14], Lighthouse are the centralized NCS, which require Fixed Landmarks for computing coordinates.

(b) Decentralized NC Systems

In decentralized systems, there is no need of such fixed nodes. Every node can act as a Landmark. The nearby nodes are used to figure out the coordinates in system. The neighbourhood nodes are found by using an active node discovery protocol. To compute coordinates, the nodes are selected by using different methods like Arbitrary chosen nodes, closely reside nodes and combination of both a hybrid approach, and are some different neighbourhood selection policies [17][21].

Various decentralized NCS are: Vivaldi [7], PIC [6], Phoenix [9] etc. In these systems the nodes compute the coordinates with reference to each other.

2.2 Models of Network Coordinate Systems

Besides this, Network coordinate systems are also classified on the basis of the model they used. They can be classified into 2 main categories: (a) Euclidean Distance model and other is (b) Matrix Factorization model [15].

(a) Euclidean Distance Based Model

For designing a NCS, the Euclidean Distance is the mostly used NC model. Within this model the H network hosts are embeds into a d-dimensional Euclidean space i.e. R^d. Let x_i as the NC of host H_i, then $x_i = (r^{i}_{1,}\ r^{i}_{2}........,\ r^{i}_{d})$ i.e. RTT for host H_i. Then the RTT among host H_i and H_j can be calculated with the help of using x_i and x_j. Hence, D^E (H_i, H_j) is defined as $D^E(H_i, H_j) = \| x_i - x_j \|$.

Although the Euclidean Distance Model is very popular among NCS but it gives raise to the problem of Triangular Inequality violation (TIV). The subsistence of TIV is usual on the internet. But it can still affect the prediction accuracy of system.

Let us consider the triangle PQR where P, Q and R are the three nodes. Now, suppose that PQ is the largest side of the triangle. If D (P, Q) > D (P, R) + D (R, Q), then PQR violates the triangular inequality property therefore called as TIV. And whichever three nodes have this TIV, they cannot be implemented in Euclidean Space, because it doesn't provide the required level of accuracy. So distance among hosts must obey the property of triangular inequality.

(b) Matrix Factorization Based Model

Matrix Factorization based technique is introduced to defeat the difficulty of TIV. The key idea behind this model is to factorize a large matrix into two smaller matrices with Singular Value Decomposition (SVD) or Non-negative Matrix Factorization (NMF). For a system with H network hosts, the $H{\times}H$ network distance matrix D can be factorized into two smaller matrices. $D \approx XY^T$, where X and Y are $H \times d$ matrices. Then $D^E = XY^T$, internet distance matrix. Then the distance between two hosts H_{i} , H_j can be calculated as:

$$\mathrm{F} = \sum_{i=1}^{N}\sum_{j=1}^{N}(D(i,j) - X_i\,.\,Y_j)^2$$

Here X_i and Y_i are the incoming and outgoing vectors respectively.

2.3 Phoenix Coordinate System Overview

Phoenix [9][15][17] is a practical dot product based, decentralized network coordinate system which makes use of matrix factorization model. Being a decentralized system, it doesn't require any fixed set of dedicated network nodes. Every node is treated evenly. Any node can calculate its coordinate with reference to any other node in network. Phoenix is used for the large scale applications so that network overhead can be distributed efficiently.

When a new node N_{new} enter in the system, it can pick n few nodes from the set S, which is the group of nodes whose coordinates have been computed and start an update method. In each round, the N_{new} measures the RTT and retrieve the incoming (X) and outgoing vectors (Y) of these nodes. The predicted distance between two nodes N_i and N_j is simply the dot product of the incoming and the outgoing vector. After that the coordinates can be determined and simplified regularly. Due to this dispersed behaviour of phoenix, any host is free to enter and leave the system. This system overcomes the problem of TIV [15].

For the early m hosts, i.e. $N \leq m$, the new node is one of the early nodes and the scale of system is very small. These early hosts communicate with each other to generate distance matrix. In this case, centralized approach is used to compute coordinates and NMF algorithm is used to get incoming and outgoing vectors. Once $N > m$, the early hosts become ordinary.

A weight based mechanism was introduced to calculate coordinate of ordinary host. The weights help to distinguish the accurate and inaccurate NCs. By doing so, the overall prediction accuracy of the system increases and minimizes the relative error.

3. COORDINATE SYSTEM ATTACKS

In coordinate systems, the attacker can attack the coordinate and the latency information. The attacker cannot always be a one who alone performs the malicious action. It can also influence others to participate or have a group of crime partners. In case of network coordinate systems, various authors discovered that an attacker can attack in two ways: The attack can be performed either by the insider, where attacker is a member of system or by the outsider, where attacker lies outside the system. Although both cases are harmful as they doesn't allow system to work properly. But the insiders' attacks are more harmful.

In above cases, the attacker attacks by reporting with false information. The attacker node can lie about information in two manners: (a) one is the fixed lies, where attacker provides same false information all the time. It is very difficult to figure out such attacks. (b) Other is continuously changing lies; in this the attacker node always provides random fake information. The system can detect this behaviour of nodes and discard the information provided by them.

Zade et. al. [20] purpose next three attacks against coordinate systems. As we previously noted, that the action of network coordinate system is depend upon the assumption that the nodes are the trusted participants in system and always report with correct information. But the attacker selected in reference set can pass the incorrect information. By blindly referring to this information a true node can calculate inaccurate coordinates. In this section, we give a review about the attacks and the key observations regarding them.

1. *Coordinate Inflation attack*: Inflation attack aims to cause valid nodes to compute large values of coordinates then the actual values. The attacker can execute this attack by making false announcement about the values or by compromising the legal nodes. The attacker node always replies with enlarge value of coordinates when a node requests for coordinates. The coordinate inflation attack is not so effective then the other attacks. It causes less damage to the system. If a valid node doesn't have any attacker as reference node, it can still compute the true value of coordinates.

2. *Coordinate Deflation attack*: Deflation attack is similar to as that of inflation attach. But this time, instead of reporting with increase value; they report with the smaller value of the coordinates. During demanding the coordinate values of nodes by other nodes, the malicious adversarial respond with the small coordinates than the original one. So the node can't move to its original position and it remains stationary in system. This can be done by minimizing the distance among the real RTT and the measured RTT. This attack prevents the victim node from being updated and remains it close to the starting point. The error remains low in this because it doesn't allow system to expand its range.

3. *Coordinate Oscillation attack*: This attack consists of both inflation and deflation of coordinate values. This attack seeks to create unsteadiness between the virtual coordinates. This results in a node to constantly change their positions in system and doesn't not result system to be in a stable state. The attacker can execute this attack by making continuous large and small random announcements about value of coordinates. Such attack aims to maximize he

relative error by reporting false mixed values. As a result, the legal nodes remain confused and deviate between the large and small values.

4. EVALUATION METHODOLOGY

In this section, we describe the methodology we use to evaluate the performance of Phoenix NC systems under these three different attacks using measured internet distance data. We also give the description of the performance metrics.

4.1 Data Set Used

To evaluate the impact of these attacks on Phoenix, we use two real time data sets. The simulation is based on both King's [7] and PlanetLab [16][12] dataset. The king is a tool which is able to obtain the network latencies between various internet hosts. It is based on direct online measurements for obtaining RTT values which are very close to the real values of RTT. It contains the RTT of 1740 DNS servers and PlanetLab contain the pair-wise latencies of about 169 internet hosts collected through the PlanetLab ping project. All these hosts are connected with each other through internet.

4.2 Experiment Methodology

To perform an experiment using data set, we consider the insider attacks. We assume that the attacks are executed by the trustful participants of the system. Those participate in computation of coordinate and have their own well defined coordinate position in network. We assume that being a part of system the attacker node have access to the information about coordinate position and the RTT of all other nodes as any other legal node have. It might happen because the attacker node has already bypassed the security mechanism like authentication or fool any legitimate node. The malicious nodes are propagated all over in the system. In this experiment, our centre of attention is toward the attacks against the performance accuracy of the coordinates systems.

Our experiment starts with running phoenix in its native form and compute relative error without introducing any malicious activity. The nodes exchange information about coordinates and RTT with each other and discover the positions in system respectively. The coordinates of nodes are believed to converge within 3 rounds. Each node select 32 nodes from its neighbourhood in its reference set randomly.

Once the system achieves the stable state, we start introducing malicious nodes. After that we compute coordinates of malicious nodes and update the position of rest of the nodes with respect to the malicious nodes. It is important for attacker to be in neighbourhood list of any node so that it can propagate incorrect information. This scenario was repeated 3 times with malicious nodes selected at randomly. We consider 30% of the total nodes as malicious. The coordinates are selected in a definite interval for all three attack scenarios.

After that we calculate and compare the relative error of phoenix under different attacks and phoenix with no attack.

4.3 Performance Metrics

The expected correctness of NCs is frequently estimated in terms of Relative Error (RE). Let A and B are the two network nodes, Da is the actual distance between these two hosts while Dp is the calculated i.e. predicted distance between A and B. Then, the Relative Error (RE) is the difference between Da, the actual distance and Dp, the predicted distance between two network host defined as:

$$Relative\ Error = \frac{|\,\text{actual} - \text{predicted}\,|}{\text{min}\ (\text{actual}, \text{predicted})}$$

The prediction accuracy of the system is higher if the RE is smaller. The RE value will be zero, if the measured distance is equal to predicted distance. RE is the main performance indicator in coordinate systems [9][12].

5. RESULTS

In our experimentation we evaluate the performance of phoenix coordinate system with and without different conditions of attacks. In this, we considered the three attacks on Phoenix system i.e. Inflation, Deflation and Oscillation attack.

5.1 Inflation Attack

We first demonstrated the inflation attack by making 30% of the total nodes as malicious. We evaluate our results on the basis that how these nodes can affect a particular victim node and cause coordinates to inflate. The total number of attacker that will affect the victim nodes are selected through the system itself as a reference nodes of that node. The size of the reference set is taken by the phoenix system itself.

It is assumed that it take about 30% of malicious nodes in every references set of nodes. In inflation attack, the attacker maximizes the RTT between the actual and the estimated time. We explicitly defined a range for this. The results are evaluated after 3 rounds of execution.

Fig.1, Fig.2 shows the effect of inflation attacks on Phoenix by using King's and PlanetLab dataset respectively. The 50 Percentile and 90 percentile of RE with king's set is increased by 14.47% and 35.08%. On the other, with PlanetLab it's raised by 23.58% and 52.33 %.

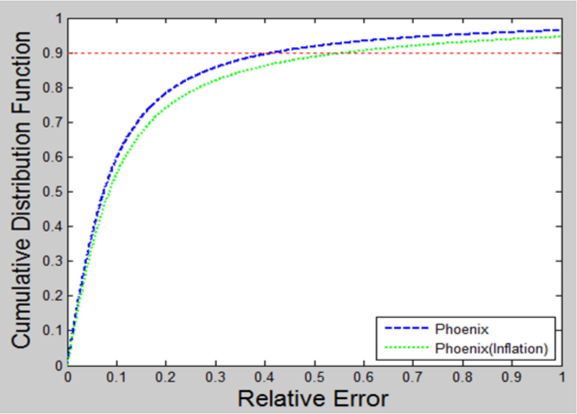

Fig. 1 Inflation attack with King's Set

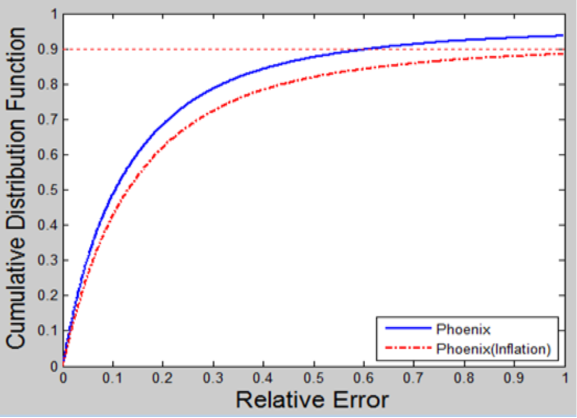

Fig. 2 Inflation attack with PlanetLab

5.2 Deflation Attack

The deflation attacks cause the nodes to report with the fake small values of the coordinates then the actual values. When a node demand for the coordinates from its neighbour nodes, the attacker present in the reference set of the nodes will announce the values smaller than the original values. They tend to minimize the RTT between the reference nodes and the target node. We consider 30% of the nodes as the attacker in whole system. The attacker nodes are spread all over. While the calculation of coordinates values attacks may be a part of the referred neighbours.

Fig. 3, Fig.4 shows the comparison of the relative error of phoenix with and without attack by using King's and PlanetLab dataset. With King's, the 50 percentile RE is increased by 11.84% and 90 percentile is 43.67%. And with PlanetLab it is 11.38% and 62.33%.

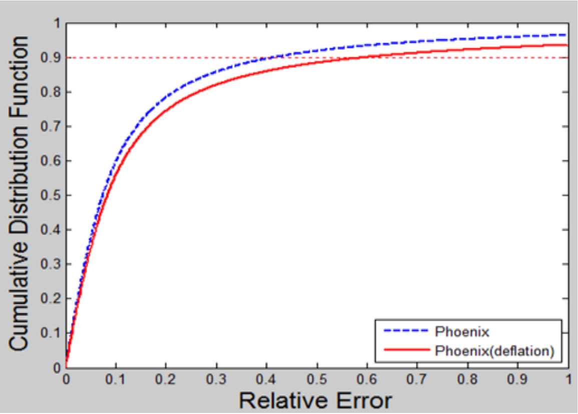

Fig. 3 Deflation attack with King's Set

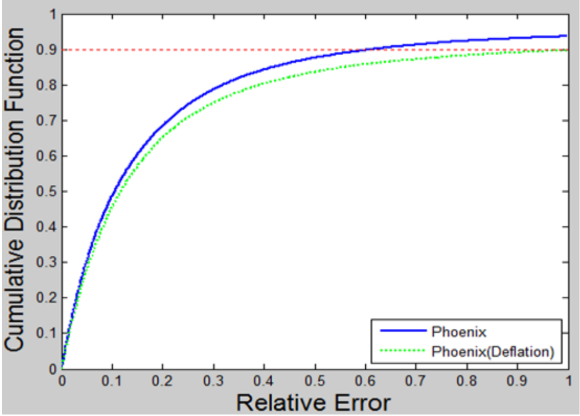

Fig. 4 Deflation attack with PlanetLab

4.2.3 Oscillation Attack

In this scenario, the attacker nodes work together to create a situation likes the disorder. It is the combination of the inflation and deflation attacks. The attacker nodes send random values of coordinates in system. The system is under random attack when some time they can increase value or they can reduce it, the system will remain in consistent state and nodes converge toward their origin of generation of system.

In Fig. 5, shows the effect of oscillation attack on Phoenix by using King's dataset. The 50 percentile RE is raised by 13.65% and 90 percentile is 39.85%.

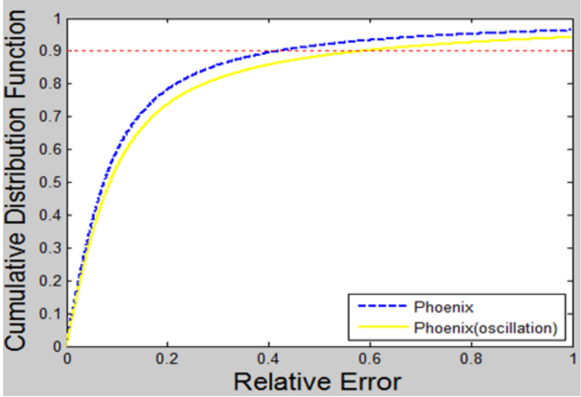

Fig. 5 Oscillation attack with King's

Fig 6 shows the effect of oscillation attack on Phoenix by using PlanetLab dataset. The 50 percentile error is increased by 12.22% and 90 percentile is 30.84%.

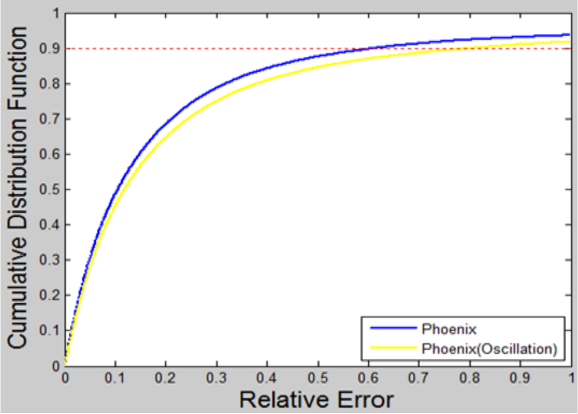

Fig. 6 Oscillation attack with PlanetLab

6. CONCLUSION

Coordinate system consists of thousand of nodes spread widely all over the internet. It is so difficult to secure such huge number of nodes. It is not difficult for any attacker to be a part of such vast system and start its action. The coordinate system should consider that any node can be unsafe node. This is important for them to not to follow the conventional cryptography. It doesn't protect them from the wrong information provided by a trustful participant.

In this paper, we have study the different attacks against the phoenix coordinate systems and also analyze the impact of these attacks on the prediction accuracy of the phoenix systems. We have found that the phoenix's inbuilt security mechanism is not secure enough to beat these attacks. In the current situations, where we are facing a number of attacks on NC, using phoenix as a system deployment is a great choice among the present ones. However, the other systems also have the security protection mechanisms with them and are protected against the malicious activities. But they don't provide the high performance rate then the phoenix. From the simulation of the experiment and the results obtained, it is clear that these attacks have an immense impact on the performance and accuracy of the Phoenix system. Because the system can't identify the random requests generated for the coordinates. From the analyses of results it is clear that it will degrade the systems performance.

7. REFERENCES

[1] Sharma, Puneet, et al. "Estimating network proximity and latency." ACM SIGCOMM Computer Communication Review 36.3 (2006): 39-50.

[2] Zhang, Rongmei, et al. "Impact of the inaccuracy of distance prediction algorithms on Internet applications-an analytical and comparative study." INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings. IEEE, 2006.

[3] Pietzuch, Peter, et al. "Network-aware operator placement for stream-processing systems." Data Engineering, 2006. ICDE'06. Proceedings of the 22nd International Conference on. IEEE, 2006.

[4] Azureus bittorrent http://azureus.sourceforge.net

[5] Zhang, Bo, et al. "Measurement based analysis, modeling, and synthesis of the internet delay space." Proceedings of the 6th ACM SIGCOMM conference on Internet measurement. ACM, 2006.

[6] Costa, Manuel, Miguel Castro, Antony Rowstron, and Peter Key. "PIC: Practical Internet coordinates for distance estimation." In Distributed Computing Systems, 2004. Proceedings. 24th International Conference on, pp. 178-187. IEEE, 2004.

[7] Dabek, Frank, Russ Cox, Frans Kaashoek, and Robert Morris. "Vivaldi: A decentralized network coordinate system." In ACM SIGCOMM Computer Communication Review, vol. 34, no. 4, pp. 15-26. ACM, 2004.

[8] Chen, Yang, Yongqiang Xiong, Xiaohui Shi, Jiwen Zhu, Beixing Deng, and Xing Li. "Pharos: accurate and decentralised network coordinate system." Communications, IET 3, no. 4 (2009): 539-548.

[9] Chen, Yang, Xiao Wang, Xiaoxiao Song, Eng Keong Lua, Cong Shi, Xiaohan Zhao, Beixing Deng, and Xing Li. "Phoenix: Towards an accurate, practical and decentralized network coordinate system." In NETWORKING 2009, pp. 313-325. Springer Berlin Heidelberg, 2009.

[10] Mao, Yun, Lawrence K. Saul, and Jonathan M. Smith. "Ides: An internet distance estimation service for large networks." Selected Areas in Communications, IEEE Journal on 24, no. 12 (2006): 2273-2284.

[11] Cox, Russ, and Frank Dabek. "Learning Euclidean coordinates for Internet hosts." (2002).

[12] Ng, TS Eugene, and Hui Zhang. "Predicting Internet network distance with coordinates-based approaches." In INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 1, pp. 170-179. IEEE, 2002.

[13] Tang, Liying, and Mark Crovella. "Virtual landmarks for the internet." In Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement, pp. 143-152. ACM, 2003.

[14] Lim, Hyuk, Jennifer C. Hou, and Chong-Ho Choi. "Constructing Internet coordinate system based on delay measurement." IEEE/ACM Transactions on Networking (TON) 13, no. 3 (2005): 513-525.

[15] Chen, Yang, Xiao Wang, Cong Shi, Eng Keong Lua, Xiaoming Fu, Beixing Deng, and Xing Li. "Phoenix: A weight-based network coordinate system using matrix factorization." Network and Service Management, IEEE Transactions on 8, no. 4 (2011): 334-347.

[16] PlanetLab, http://www.planet-lab.org/

[17] Wang, Gang, Shining Wu, Guodong Wang, Beixing Deng, and Xing Li. "Experimental study on neighbor selection policy for Phoenix Network Coordinate system." In Ultra Modern Telecommunications & Workshops, 2009. ICUMT'09. International Conference on, pp. 1-5. IEEE, 2009.

[18] Kaafar, Mohamed Ali, Laurent Mathy, Chadi Barakat, Kave Salamatian, Thierry Turletti, and Walid Dabbous. "Securing internet coordinate embedding systems." ACM SIGCOMM Computer Communication Review 37, no. 4 (2007): 61-72.

[19] Kaur, Manpreet, and Akhil Sharma. "Attacks on Phoenix Coordinate System."

[20] Zage, David, and Cristina Nita-Rotaru. "Robust decentralized virtual coordinate systems in adversarial environments." ACM Transactions on Information and System Security (TISSEC) 13, no. 4 (2010): 38.

[21] Zage, David John, and Cristina Nita-Rotaru. "On the accuracy of decentralized virtual coordinate systems in adversarial networks." In Proceedings of the 14th ACM conference on Computer and communications security, pp. 214-224. ACM, 2007.