IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 2, Issue 1, Feb-Mar, 2014
**ISSN: 2320 - 8791**
**www.ijreat.org**

# Multi-Level Security and Detection against Clone Attack in Military Scenario

**A.Madhumitha[1], Dr.B.Amutha[2]**

*[1]Department of Computer Science and Engineering, SRM University, Tamil Nadu, India- 603203*
*[2]Professor, Department of Computer Science and Engineering, SRM University, Tamil Nadu.*

## Abstract

*Wireless Sensor Networks are vulnerable to the clone attack. The three levels of security clearance in military are Topsecret, Secret and Confidential, which are used as part of a method to control access to information that should not be freely available to all personnel. To avoid disclosing the information of the clearance level by clone attack, Randomly Directed Exploration Protocol will be implemented which is location based system identification, where every location will have a group leader. The Group leader will generate a random number with time stamp to the available system. The chaotic code technique is generated to encrypt the random number, timestamp, message and the ID's of the system. A witness system is to detect the cloned system and will perform random key prediction, ID based prediction and secret key encryption to the clearance levels. Henceforth the level of Security is conceptually increased by using chaotic code technique.*

**Keywords: Randomly Directed Exploration, chaotic code, Random number, Time Stamp**

## 1. Introduction

A wireless sensor network (WSN) is a collection of nodes organized into a cooperative network. Each node consists of processing capability, which is one or more micro controllers, CPUs or DSP chips. Each node may also contain multiple types of memory that is program, data and flash memory and also have a RF Transceiver.

Nodes can sense the environment and communicate the information gathered from the monitored field through wireless links and the data is forwarded, possibly via multiple hops relaying, to a sink that can use it locally, or is connected to other networks (e.g. the Internet) through a gateway.

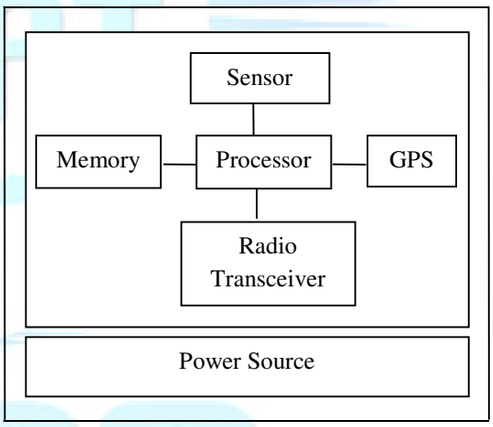


**Figure 1.**Basic Components of WSN

Wireless Sensor Network challenges are energy efficiency, robustness, scalability, systematic design, privacy and security. WSN have gained a great deal of attention in the past decade due to their wide range of application areas and formidable design challenges. The constraints are resource and design constraints. The resource constraint has low bandwidth, limited storage, limited processing and limited energy whereas the design constraint has application and environment dependent. It consists of hundreds and thousands of low-cost, distributed sensor nodes, which scatter in the surveillance area randomly, without attendance [12]. If the operation environment is hostile, security mechanisms against adversaries should be taken into consideration. Some of the security threats of

WSN are Sybil attack, Wormhole attack, Denial of Service attack. Among many physical attacks the node clone is a serious and dangerous one. WSN are exposed to the clone attack. Clone attack is defined as node replication attack. The cloned nodes that seem legitimate can freely join the sensor network and then significantly enlarge the adversary's capacities to manipulate the network maliciously and fetch the information. For example, those cloned nodes occupy a considerable position in the network and cooperatively corrupt the collected information. With a large number of cloned nodes under command, the attackers gain access to the whole network. Furthermore, the node clone will aggravate most of inside attacks against the sensor networks. Applications of WSN are

- Monitoring groundwater contamination and Ecosystem.
- Civil Structural, Health Monitoring.
- Rapid Emergency Response
- Military

In Military application, all classified information is divided into one of three categories they are Topsecret, Secret, Confidential clearance level [9]. Topsecret is applied to information that reasonably could be expected to cause exceptionally grave damage to the national security if disclosed to unauthorized sources. Secret is applied to information that reasonably could be expected to cause serious damage to the national security if disclosed to unauthorized sources. Confidential is applied to information that reasonably could be expected to cause damage to the national security if disclosed to unauthorized sources.

## 2. Analysis of Clone Attack

In the wireless sensor networks, due to the low security mechanism, the attackers can easily attack the WSN and perform many illegal activities. Tamper resistant hardware is expensive, so most wireless sensor networks are composed of unshielded sensor nodes. An adversary can easily attack, analyze and clone the unshielded sensor nodes and create replicas and insert them into the network.

The two novel node clone detection protocols with different trade-offs on network conditions and performance [12]. The first one is based on a Distributed Hash Table (DHT) in which Chord algorithm is used to detect the cloned node, every node is assigned with the unique key, and before it transmits the data it has to give its key which would be verified by the witness node. If same key is given by another Node then the witness node identifies the cloned Node. The second one is based on the Distributed Detection Protocol named randomly directed exploration perform good communication performance in dense sensor network, which is same as DHT, but it is easy and cheaper implementation. Here every node only needs to know the neighbour-list containing all neighbour IDs and its locations. So that can detect node clone with high security level and holds strong resistance against adversary's attacks. No any specific routing protocols or infrastructures are demanded in the RDE protocol [2].

Randomly Directed Exploration protocol is a location based nodes identification, where every region or location will have a group leader. The Group leader will generate a random number with time stamp to the available nodes in that location. A witness system is generated to verify the cloned system's activities and it will act in between the source and destination system. The witness system will perform random key prediction and ID based prediction in Topsecret level. In the secret level, random key prediction security scheme is used for detection of replication attack. In the Confidential level, the Secret key encryption will be performed for security purpose. The chaotic technique is generated to encrypt the ID's of the system or node, random number, timestamp and also the data is encrypt with the Data Encryption Standard algorithm to enhance the security.

## 3. Algorithm and Methodology

The distributed detection protocol is to make use of the DHT mechanism to form a decentralized caching and checking system that can effectively detect cloned nodes.

### 3.1. Chord Algorithm

In the Peer-to-Peer system, Chord algorithm is used to detect the cloned node. P2P systems are popular due to low startup cost, high scalability at very low cost. The p2p system offers a way to aggregate and make use of tremendous computation and storage resources on computer across the internet. Chord Algorithm is a peer-to-peer lookup service [3]. It solves problem of locating a data item in a collection of distributed nodes. The Chord characteristics are simplicity, provable correctness, and provable performance. The chord maintains routing information as nodes join and leave the system. Given a key (data item) it maps the key onto a node (peer).

The consistent hash function assigns each node and each key an m-bit identifier using SHA 1 (Secure Hash Standard).

m = any number big enough to make collisions improbable
Key identifier = SHA-1(key)
Node identifier = SHA-1(IP address)

Identifiers are arranged on a identifier circle $2^m$ called the Chord ring (Shown in Figure 4.1) A key k is assigned to the node whose identifier is equal to or greater than the key‘s identifier. This node is called successor(k) and is the first node clockwise from k [6].

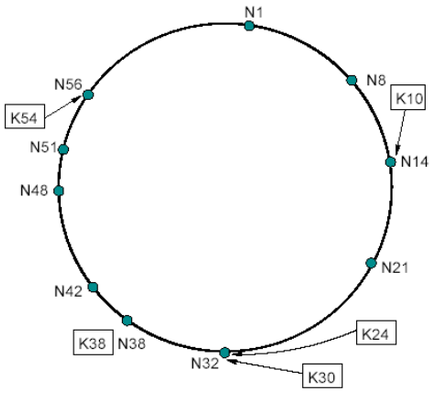


**Figure 2.**Chord Ring

The finger table allow faster location by providing additinal routing information than the successor node. i is the finger table index.

**Table 1.**Notation for chord

| Notation | Definition |
|---|---|
| Finger[i].start | $(n + 2^{i-1})\bmod 2^m$ 1=<i=<m |
| .interval | [Finger[i].start, Finger[i+1].start) |
| .node | First node>=n. Finger[i].start |
| successor | Next node on the identifier circle finger [1].node |
| predecessor | Previous Node on the identifier circle |

To Find Successor in Chord
1: n.findSuccessor(id) begin
2: n’← findPredecessor(id)
return n’.successor(id)
3: end
4: n.findPredecessor(id) begin
5: n' ← n
while id ∉ (n',n'.successor()) do
6: n’ ← n’.closestPrecedingFinger(id)
7: end
8: end
closestPrecedingFinger(id)
1: n.closestPrecedingFinger(id) begin
2: for i ← m to 1 do
3: if finger[i].node ∈ (n; id) then
4: return finger[i]:node
5: end
6: end
7: return n
8: end

The important characteristics are

- Each node stores information about only a small number of nodes (m)
- Each nodes knows more about nodes closely following it than about nodes farer away
- A finger table generally does not contain enough information to directly determine the successor of an arbitrary key k

Each node n contains a routing table with up to m entries where m is number of bits of the identifiers.

Finger Table(S)= successor $(n + 2^{i-1})$

### 3.2. Stemming Algorithm

Stemming is the process for reducing the inflected or sometimes derived words to their stem, base or root form which is generally a written word form. The process of stemming is often called conflation. These programs are commonly referred to as stemming algorithms or stemmers. The applications are Information retrieval and also ussage in commercial products [8].

The matching algorithm uses a stem database. An example, a set of documents that contain stem words. These stems are not necessarily valid words themselves but rather common sub-strings, as like the "brows" in "browse" and in "browsing". In order to stem a word the algorithm tries to match it with stems from the database, applying various constraints, such as on the relative length of the candidate stem within the word, let us consider one more example the short prefix "be", which is the stem of such words as "be", "been" and "being", would not be considered as the stem of the word "beside". By analysing various sources of information from the internet the below set of keyword or term are mentioned for each clearance level (refer Table 2).

**Table 2.**Terms for Clearance Level

| Clearance level | Term/Keywords |
|---|---|

| | |
|---|---|
| Topsecret Level | Missile, Launch, Uranium, Fusion, Torpedoes, Attack chemical, Weapon, Coded-name, Activate, Reactor, Ballistic, Designator, Guided, Warheads, Sidewinder, Bio-gas Affect, Electromagnetic, Plutonium |
| Secret Level | Arm, Buckshot, Rifle, Supply, Power, Explosive, Calibre, Range, Rounds Speed, Operations, Defence, Manufacture, Design, Versions, Launcher, Maintenance, Produce, Aircraft, Tanker, Submarine |
| Confidential level | Generals, Soldiers, Register, Number, Grade, Navy, Designation, Batch, Address, Responsibilities, Age, Date, Contact, Service, Commander, Air-force, Details, Conscripts, Group, Field, Battalion, Records |

The theoretical stemmer calculation is calculated by "Number of words per conflation class". The average size of the groups of words converted to a particular stem (regardless of whether they are all correct). Thus, if the words "engineer", "engineering" and "engineered" were all stemmed to the stem "engineer", then the size of that conflation class would be 3. If the conflation of 1,000 different words resulted in 250 distinct stems, then the mean number of words per conflation class would be 4 [14].

The metric is obviously dependent on the number of words processed, but for a word collection of given size, a higher value indicates a heavier stemmer. The value is easily calculated as follows,

WC = Mean number of words per conflation class
N = Number of unique words before Stemming
S = Number of unique stems after Stemming

$$\mathrm{WMC} = \frac{\mathrm{N}}{\mathrm{S}}$$

The Theoretical Calculation to find the Threshold value is calculated by Term Frequency. The tf-idf stands for term frequency-inverse document frequency and the tf-idf weight is a weight often used in information retrieval and text mining [5]. It is a statistical measure used to evaluate how important a word is to a document in a collection or corpus. The importance increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word in the corpus. For a given corpus D, and then the tf-idf is then defined as

$$(\mathrm{tf} - \mathrm{idf})_{i;j} = \mathrm{tf}_{i;j} \times \mathrm{idf}_i$$

Where, Term Frequency - $tf_{i;j}$ and inverse document frequency- $idf_i$.

In the tf-idf, the Term Frequency (tf) alone is considered to categorize the documents to the level of clearances**,** tf which measures how frequently a term occurs in a document. Since every document is different in length, it is possible that a term would appear much more times in long documents than shorter ones. Thus, the term frequency is often divided by the document length as a way of normalization.

$$\mathrm{tf(t)} = \frac{(\text{No.of times term t appears in a document})}{(\text{Total no.of terms in the document})}.$$

The inverse document frequency (idf) is explained to understand the concept of td-idf, it is obtained by dividing the total number of documents by the number of documents containing the term $t_i$, and then taking the logarithm of that quotient.

$$\mathrm{idf}_i = \log \frac{|D|}{|\{d : ti \in d\}|}$$

Where |D|: total number of documents in the collection | {d: ti ε d}|: number of documents where the term $t_i$ appears. To avoid divide-by-zero, can use 1 + | {d: $t_i$ ∈ d}|.

The Theoretical calculation for tf are as follows, Determine the set of terms $t_i$ for each level of clearance.The term frequency (tf) for a given term $t_i$ within a particular document $d_j$ is defined as the number of occurrences of that term in the $d_j$th document, which is equal to $n_{i;j}$ : the number of occurrences of the term ti in the document $d_j$.

$$\mathrm{tf}_{i;j} = n_{i;j}$$

The term frequency is often normalized to prevent a bias towards larger documents, as shown below,

$$\mathrm{tf}_{i;j} = \frac{n_{i;j}}{\sum_k n_{k;j}}$$

Where $n_{i;j}$ is the number of occurrences of the term $t_i$ in the document $d_j$. Note that we are using the total number of terms for normalization. Instead we can use the maximum as well [5].

Example 1: Consider a document containing 1000 words wherein the word nuclear appears 3 times. The

term frequency (i.e., tf) for nuclear is then (3 / 100) = 0.03. Since the term appears are categorized under Topsecret level of clearance the priority given to Topsecret level.

### 3.3. Chaotic Code Technique

The chaotic code methodology is proposed so that the clone node cannot presumption the ID's and also the random number with the timestamp. It involves the following steps

Step 1: Source ID: The user's source id is 8 digit decimal numbers.

i. Convert the decimal number to binary value by splitting the number into 4 digits.
ii. The first 4 digit corresponding binary value is taken and number it from right to left as (1 to n) values.
iii. Arrange the binary value based on the even and odd position respectively from right to left. Name the even position binary value's as A and odd position as B.
iv. Perform XOR operation for A and B , by making the first term of A and B as a input , Similarly Second term of A and B, Third term of A and B so on from right to left direction. Add '0' as MSB when the length of binary value doesn't match.
v. Reduce the Binary value to 4 bit by performing OR operation that will be the final binary values of first 4 digit decimal number
vi. Similarly for the next 4 digit decimal value perform the above operation (2-5) and obtain the result.
vii. The result of these two operations is combined and make it as 8 bit of binary, if the LSB of the 8 bit binary value is '1', add '+1' and obtain the final 8 bit binary value for the source id.

Step 2: Destination ID: The Destination id is 8 digit decimal numbers.

i. Similar process as Source id process but NAND operation is used instead of XOR operation.
ii. If the LSB of the combined result of 8 bit binary value is '0' add '+1' and obtain the final 8 bit binary value for the destination id.

Step 3: Group ID: The group id is 8 digit decimal numbers.

i. Similar process as Source id process but NOR operation is used instead of XOR operation.
ii. If the MSB of the combined result of 8 bit binary value is '1'add '+1' and obtain the final 8 bit binary value for the group id.

Step 4: Random Number: The random number is 4 digit decimal numbers.

i. From right to left take the first and last digit and perform the binary conversion which is the first set result. Similarly take the second and third digit and perform the binary conversion which is the second set result.
ii. The result of first set binary value is partition for every 2bits from right to left and perform OR operation within the corresponding 2bits. Similarly for second set result also the same operation is done. Add '0' as MSB when the bit partition length doesn't match
iii. The Processed result of the first and second set is combined and made as 8 bit binary value.

Step 5: Timestamp: The timestamp is classified based on 12hrs and 24hrs.
For 12hrs the time is mentioned with 0 as starting e.g. 04:59

i. Perform binary conversion for the hour hand and minute hand time separately.
ii. For each result perform the AND operation first and NOT operation
iii. Add '1' only when the MSB is '0' for the binary value which got after the NOT operation.

Step 6: Predecessor ID and Successor ID: The predecessor id and successor id is 8 digit decimal numbers.

i. Divide the 8 digit decimal number into 2 digits and perform the binary conversion for each 2 digit.
ii. Perform NOR operation for the each binary results. From right to left Take the first and last term, Perform NOR operation. Similarly take the second and last before term, Perform NOR operation. Similarly for the rest of terms in the binary value.
iii. Combines the first 2 set of binary value and do NOR operation similarly for next set, Repeat the process till 8 binary bits is obtained.

STEP 7: To the data, Data Encryption Standard algorithm is applied. The DES is a symmetric-key block cipher. It operates on plaintext blocks of a given size (64-bits) and returns cipher text blocks of the same size. DES results in a permutation among the 2^64 possible arrangements of 64 bits, each of which may be either 0 or 1. Each block of 64 bits is divided into two blocks of 32 bits each, a left half block L and a right half R [11]. The DES Decryption uses the same algorithm as encryption, except that the application of sub keys is reversed [11]. In DES Encryption there will be 16 rounds of function.

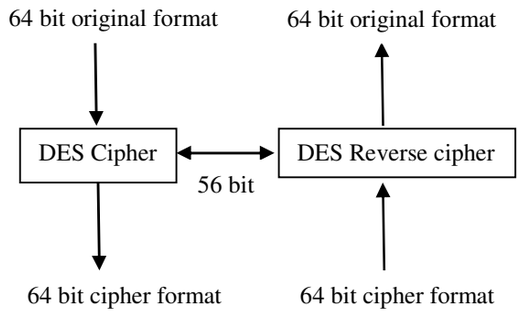


**Figure 3.**Encryption and Decryption of 64bit in DES

## 3.4. Methodology

### 3.4.1. Network Construction

Dynamic network is created. In the network, there are Group Leader Node, N number of Nodes, Witness Node and Destination Node. In a network, nodes are interconnected with the admin, which is monitoring all the other nodes. All nodes are sharing their information with each other's.

### 3.4.2. Randomly Directed Exploration Protocol

Apply chord algorithm, to make the nodes in chord ring form. In chord algorithm, the Neighbour nodes information of the requested node is verified. The verified Id's and locations of the nodes are used to detect the clone node. For this purpose, create the list of the neighbour nodes information for each node so that the Server or Witness node can verify the nodes request [2][12].

### 3.4.3. Data Set Classification

Source node is sign in with the user name and password. Apply stemming algorithm to filter the matched keywords in the data set and calculate the threshold value which has seven possible ways (refer Table 3).

Based on the threshold values the results are transmitted to the node. Also based on the query and result the security information will be asked by the network during communication. If the result contains top secret clearance type of data then nodes have to answer all the security information. If the results contain secret clearance data, then nodes have to answer all the security information except the predecessor and successor information. If the results contain only confidential clearance results, then the nodes have to answer only the Encryption and decryption information.

**Table 3.**Priority and Security

| Level of clearance | Priority | Security |
|---|---|---|
| Topsecret only | Topsecret level of clearance | All the security mechanism is provided (Source id + destination id +Group ID + Random no + Time stamp + Pred. id + Successor id + Data Encrypted by DES) |
| Topsecret and secret level | | |
| Topsecret and confidential level | | |
| Topsecret, Secret and confidential level. | | |
| Secret only | Secret level of clearance | Except unique key generation |
| Secret level and confidential level | | |
| Confidential level | Confidential level of clearance | Secret key encryption |

### 3.4.4. Random Key Pre-distribution

Random Key pre-distribution [13] security scheme is implemented in the sensor network. That is, each node is assigned a number randomly with time stamp from the group leader in the source group. Then the group leader will transmit random number which was generated with respect to that time Stamp to the Witness node [1]. Witness node will now check the random number which is generated with the user information. If both the informations are matched then the Witness node will confirm that the node is genuine one.

### 3.4.5. Witness Node Verification for clearance level

When the data is categorized into Topsecret clearance level means the priority is given to Topsecret level and the following operation should be perform. The Data along with ID's of the system, random number and time stamp undergo chaotic technique in the witness node. To transfer data to the destination first of all the witness node request the group leader for the random number and timestamp once it's get verified by the witness node, Again Witness node check for the predecessor and successor id if it's again

verified as genuine only the witness node send the data to the destination node otherwise witness node declare it as cloned node and terminate the process.

Similarly for Secret clearance level when it the priority is given to secret level and the following operation should be perform. The Data along with ID's of the system, random number and time stamp undergo chaotic technique in the witness node. To transfer data to the destination the witness node request the group leader for the random number and timestamp once it's get verified by the witness node as genuine it send the data to the destination node.

When the data is categorized into confidential clearance level means perform Secret key encryption in the witness node and once verified as genuine node it send the data to the destination node.

## 4. Overview

The Security system is designed in such a way to prevent the system or node from clone attack and also to secure the data or document while transferring to the destination system in the Military application of Wireless Sensor Network. Consider that the data set contains missile information and missile launch information, and then it is considered as Topsecret clearance level data. If the data set contains the operation underlying details and military equipment then it is considered as Secret clearance level data. If the data set contains army member's details, then it is considered as confidential clearance level data. When the source system tries to send the data to the destination; the witness system/node act in between them will verify whether the source is genuine or not.

In witness system, Apply stemming algorithm to the data, filter and extract the matched term (keyword) in the data set and calculate the threshold value. Based on the threshold values the data is categorized into the clearance level. The security is provided based on the priority to the three clearance level. Only the witness node confirms the sender node, the data is send to the destination, which once the sender is verified as genuine. If node specified information and the internal information are varied then the witness node will identify that cloning or some mal practice has occurred and the packets are discarded by the witness node.

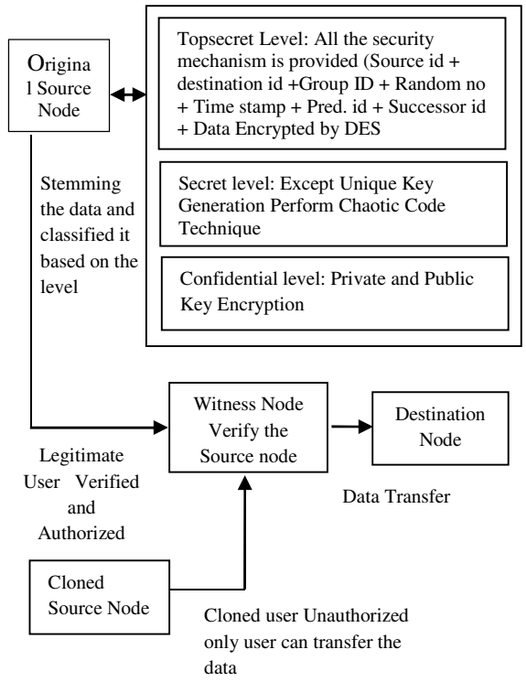


**Figure 4.**Architecture

## 5. Conclusion

In Military the level of clearance control information are transferred safely to the destination by using chaotic code technique. The Randomly Directed Exploration and DHT based protocol prevent the system from clone attack. The memory requirement of the protocol is almost optimal and the protocols achieve high detection probability. The witness node can increase the detection rate of the Cloned Nodes by verifying the time Stamp and random number.

## 6. Reference

[1] Amutha.B, V.Nivedha Devi," AMNI'09 PROTOCOL," International Journal of Information Technology and Knowledge Management, Volume 2, Number 2, Publisher page 297-303, July-December 2009.

[2] Zhijun Li, Member, IEEE, and Guang Gong, Senior Member, IEEE, "Randomly directed exploration: An efficient node clone detection protocol in wireless sensor networks" in proceeding Mobile Adhoc and Sensor Systems, 2009. MASS '09. IEEE 6th International Conference, Publisher Page 1030-1035, October 2009

[3] Balakrishnan.H, M. F. Kaashoek, D. Karger, R. Morris, and I. Stoica, "Looking up data in P2P systems," Communication ACM, volume 46, number 2, Publisher page 43–48, 2003.

[4] Brooks.R, P. Y. Govindaraju, M. Pirretti, N. Vijaykrishnan, and M.T. Kandemir, "On the detection of clones in sensor networks using random key pre-distribution," IEEE Transaction Systems, CYBERNETICS, Revised, volume 37, number 6, Publisher page 1246–1258, November 2007.

[5] Term frequency and inverse term frequency: http://www.tfidf.com/

[6] Liben-Nowell.D, Stoica.I, R.Morris, D. R. Karger,M. F.Kaashoek, F.Dabek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup protocol for internet application applications," IEEE/ACM Transaction Networking, volume 11, number 1, Publisher page 17–32, February 2003.

[7] Parno.B, A.Perrig, and V.Gligor, "Distributed detection of node replication attacks in sensor networks," Proceeding IEEE Symposium Security Privacy, Publisher page 49–63, 2005.

[8] Stemming: http://xapian.org/docs/stemming.html

[9] The three clearance level:

http://govcentral.monster.com/security-clearance-jobs/articles/2330-3-levels-of-security-clearance

[10] The phishing attack:

http://computer.financialexpress.com/sections/news-analysis/1194-drdo-s-zero-day

[11] DES Algorithm: cs.ucsb.edu/~koc/cs178/docs/04-des/chap07.pdf.

[12] Zhijun Li, Member, IEEE, and Guang Gong, Senior Member, IEEE, "On the Node Clone Detection in Wireless Sensor Networks," IEEE/ACM transactions on Networking, 2013.

[13] Random Key Pre distribution Schemes for Sensor Networks:
www.cs.cmu.edu/~haowen/chan_randomkey.pdf.

[14]Stemmingcalculation:
http://www.comp.lancs.ac.uk/computing/research/stemming/general/perfomance.htm