Serdica J. Computing **14** (2020), No 1–2, 51–63

Serdica
Journal of Computing

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

COMMON VULNERABILITIES AND EXPOSURES FOR EMAILS

Falak Ussien Hasan

ABSTRACT. Ontologies are used for knowledge management, concept definition and semantic search. Ontology is applicable to the organization and management of knowledge in a single given domain. This paper is an attempt to construct a knowledge base for email vulnerabilities using ontologies. It presents a method of building relations among CVEs email entries as established by the MITRE Corporation's weaknesses data base. The use of ontology is illustrated with queries analyzing software products from a security manager's point of view. This work is based on the MITRE community effort CWE List, Version 3.1—Research concept view CWE-1000.

1. Introduction. The use of emails has become something very common among people. We can see emails being used everywhere around the world when it comes to exchanging data. With billions of messages sent every

ACM Computing Classification System (1998): I.2.4, K.4.1, H.4.3.

Key words: mail weaknesses, ontology concepts, ontology principles, weaknesses, weakness concepts.

This paper is partly supported by the National Scientific Program "Information and Communication Technology for a Single Digital Market in Science, Education and Security (ICTinSES)", financed by the Ministry of Education and Science in Bulgaria. This article presents principal results of the author's doctoral thesis "Supporting CWEs and CVEs Knowledge Base for Mailing Aspects".

day, email has become one of the most up-to-date present electronic technologies in the recent 40 years [2]. The importance of the presented effort can be seen in the current context of the cybercrime as described below.

Since 2013, criminals from the Carbanak cyber gang have attempted to attack up to 100 banks, e-payment systems and other financial foundations in around 30 countries [9]. The attacks remain active. The gang has been able to steal approximately one billion USD from financial foundations worldwide. Hackers have penetrated the financial systems of several banks in countries worldwide, including Russia, Japan, Switzerland and the United States.

In August 2016, a hacker published the phone numbers and e-mail addresses of 200 former and current Democratic Party members of the US Congress. The Wall Street Journal confirmed that hackers had given America's enemies important data, such as personal information about members of the Intelligence, Military Services and Foreign Relations Committee of the House of Representatives.

Confidentiality of information plays a significant role even in the elections of the largest countries. This appeared very clear in the 2017 presidential elections in the United States. Reports showed that the US agencies had been able to identify the hackers responsible for e-mails piracy ahead of the elections. Russian hackers were alleged to be involved in the incident [6].

Malicious domains mimicking legitimate political websites were discovered and shut down. Russia-linked accounts had been used by third parties to purchase social media ads. Social media nowadays play a sustainable role in election campaigns. Facebook, for instance, has set up a war room to tackle election interference. Twitter, on the other hand, has removed more than 10,000 bots posting messages that urge people not to vote [7].

The authors of [1, 13] show that women are more likely to fall as victims of phishing than men. Young people between 18 and 25 years old get into similar situations. This may be due to lack of awareness about phishing threats. According to RSA, many online frauds have been identified as phishing attacks, and they have increased vigorously over the years.

In this paper, the researcher has made a comprehensive review of the threats and vulnerabilities with the intention of highlighting the existing security-related shortcomings of email systems. The aim of the research is to collect and organize the email vulnerabilities for the end user's use. The security manager's perspective has been taken into consideration. This research is a step towards email security protection.

This paper also aims at presenting the approach that is being used concerning email weaknesses and vulnerabilities from the software security

manager's point of view. The cyber security aspects of mailing security process of email software products are an example of this. The approach consists of:

1. The construction of a suitable knowledge base.
2. A declaration of the interactive process of the investigation of the related weaknesses and vulnerabilities.
3. The weakness and vulnerabilities, represented in an ontology.
4. SPARQL queries used to extract the necessary information from the ontology.

The ontology includes knowledge for more than 350 CWE entries with 1150 CVE entries without duplications. Concerning the entries with duplications, the ontology includes knowledge for more than 450 CWE entries with 1590 CVE entries. In both cases, each individual has nine data properties and several object properties representing different types of relationships among the individuals depending on their abstraction level.

2. The Knowledge Base. The knowledge base subject of this research represents the knowledge extracted from the CWE data base. The last one is drafted in short for presentation purposes in the next section. The knowledge base elements are discussed in the following sections.

2.1. CWE Entries and Weaknesses. Software weaknesses are flaws, faults, bugs, vulnerabilities and other errors in software implementation, code, design, or architecture. In the event of their being left unaddressed, they could result in systems and networks being vulnerable to attack [14]. The community in MITRE Corporation has compiled a list of the common weaknesses and named it Common Weakness Enumerated (CWE). The list is a kind of dictionary that can be used as a common language and standard reference for use by developers, researchers, security managers and vendors. The number of email entries detected in this research in CWEs database is around 350, including the intermediate nodes. The entries are organized in the form of a tree according to the abstraction relations between them.

The weakness entries are represented in CWE node comprises of the weakness identity ID, Name, Abstraction, Description, Extended Description, Relationships, Common Consequences, Likelihood of Exploit, Demonstrative Examples, Observed Examples and others.

2.2. CVE Entries and Vulnerabilities. Vulnerability represents a mistake in the software that can directly be used by a hacker to gain access to

a system or network [4]. CVE is recognized to represent one identifier and standardized description for each vulnerability or exposure, and accepted as one sharable language between various audiences [5]. Each CVE entry has three main properties. The first property is CVE identifier (CVE-ID) that represents the unique specific number of the entry. The second one is "Description" that provides a full definition and information about the entry. The third property is "References" that represents the recorded and deployed resources about the vulnerability exposure and exploration. This work essentially concentrates on the analysis of vulnerability concepts considering the semantic ontology concepts to construct the knowledge base. In addition to the properties CVE-ID, Description and References, there are other properties, namely Component, ProductVendor, Version, Attacker, RootCause, Impact, and Vector. These properties have been extracted and analyzed from the property Description and included in this work. This research concentrates on email weaknesses and vulnerabilities. The ontology of these entries is represented in the following sections.

2.3. Classes. They are the main elements of the ontology components and can be interpreted as sets of individuals [8], exhibiting a concrete representation of concepts. The classes in our ontology are as follow:

- CVE Entries. This class contains all CVE entry individuals. The individuals in this class have CVE-ID, Description, Component, ProductVendor, Version, Attacker, RootCause, Impact, and Vector data type properties.
- CWE Entries. This class contains all CWE entry individuals from the target domain. It has a number of subclasses, namely Classes, Bases, Variants, and Views. All individuals in this class have ID and Name data type properties.
- Classes – this subclass contains all the instances in CWE class abstraction level.
- Bases – this subclass contains all the instances in CWE base abstraction level.
- Variants – this subclass contains all the instances in CWE variant abstraction level.
- Views – this subclass contains all the instances in CWE view abstraction level.

2.3. Properties. They represent the main basic elements of ontology components. Properties have binary relations on the individuals [12] to describe the individual's membership in a certain class through constructing some restrictions. Furthermore, they exhibit a concrete representation of concepts/instances. Properties in work ontology are defined as the following:

2.3.1. Object Properties. They represent relationships between classes (individuals). The object properties in our ontology are:

- Vuln Type – this property links a CVE individual to a CWE individual.
- Has Member – this property links an individual from the Views class to an individual from the Classes class. The inverse of this object property is "Member Of".
- Has Child – this property links an individual from the Classes class to an individual from the Classes, Bases, or Variants class. The inverse property of this object property is "Has Parent".
- Has Base Child – this property links an individual from the Bases class to an individual from the Bases, or Variant class. The inverse property of this object property is "Has Base Parent".
- Has Variant Child – this property links an individual from the Variants class to an individual from the Variant class. The inverse property of this object property is "Has Variant Parent".

2.3.2. Data Type Properties. Data properties link an individual to an XML Schema Data type value or an RDF literal [11]. They link an individual from a class to a data type (integer, string, Boolean etc.). The data type properties in our ontology are:

- ID – this data type property represents a CWE entry identifier. It is an integer and is specified for all CWE entries.
- Name – represents a CWE entry name. It is a string and is specified for all CWE entries.
- Description – this data type property represents CWE entry description. It provides short explanations and definitions about Classes, Bases, and Variants entries. It is a string and is specified for all Classes, Bases and Variants class CWE entries.

- Extended Description – this data type property represents an extended description of CWE entry that contains additional information about the entry. It is a string and is optionally specified for Classes, Bases and Variants class entries.
- Objective – this data type property declares the View CWE entry objective explanation. It is a string that explains the view approach aim.

The CVE Entries has the following data properties:

- CVEID – this data type property is the vulnerability CVE entry identifier.
- CVEDescription – this data type property is the CVE entry definition and explanation.
- ProductVendor – this data type property is the name of the product and/or vendor.
- Component. This data type property is the name of the component part of the software product.
- Version. This property is the name of the version(s) of the product.
- Attacker. This data type property is the name of the attacker property.
- RootCause. This data type property is the name of the Root Cause property. It contains any additional information about the techniques and tools for the CVE entry.
- Impact. This data type property is the name of the impact property. It contains information about the effects of the software product.
- Vector. This data type property is the name of the vector property. It contains information about the tools and techniques used in attacking.

The property characteristics are as follows:

- Functionality – the functionality characteristic is a property that can have only one (unique) value for each instance [10]. It is defined for ID, Name, Description, and all object properties.
- Transitivity – in the event that a property P is a transitive property; if the pair (x, y) is an instance of P, and the pair (y, z) is also an instance of P, we can infer that the pair (x, z) is consequently an instance of P [10]. It is defined for Has Member, Has Child, Has Base Child and Has Variant Child properties.

2.4. Individuals. They are the other main ontology components. They represent objects in the domain.

An individual is created with a proper name (an instance CWE-20). Its type class can be Views, Classes, Bases, or Variants.

The individuals are related through the properties HasChild, HasBaseChild, HasVariantChild or HasMember.

3. Navigation. Concepts are organized in triple patterns (subject, predicate, object) [3]. SPARQL statements retrieve the knowledge concepts.

The following queries (cases) are based on a security system for email vulnerabilities according to the security manager's perspective as shown in Figure 1. For CWE Entries and CVE Entries in Table 1, 2 and 3 see [14, 5]. View the List of Weaknesses: By Research Concepts.

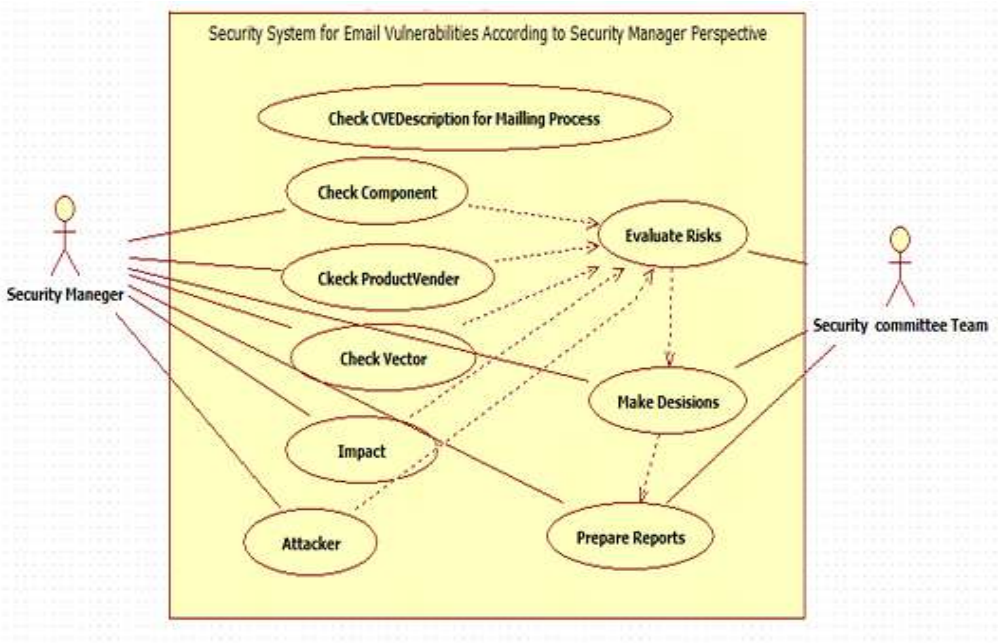

Fig. 1. The interaction of vulnerability concepts according to the security manager's perspective

CASE 1: Demonstration obtaining all email CVE entries related to mail processing:

```
PREFIX : <http://www.semanticweb.org/falak/ontologies/2018/0/20/untitled-ontology-337#>
        SELECT DISTINCT *
        WHERE { ?CVEID  :CVEDescription ?CVEDescription;
                        :Component ?Component;
```

```
                                :ProductVendor ?ProductVendor;
                                :Version ?Version;
                                :RootCause ?RootCause;
                                :Impact ?Impact;
                                :Vector ?Vector;
                                :VulnType ?VulnType.
                    ?VulnType :Name ?CWEName;
                                :Description ?CWEDescription.
filter( regex(?CVEDescription, "mail")) } ORDER BY  ?VulnType
```

This query shows vulnerabilities CVEs and their CWE entry(s) with properties Components, ProductVendor, Version, Root Cause, Impact, and Vector, and it declares the CWE entry properties ID, Name, and Description, the query implemented on the basis of the "mail" keyword using the CVE Description property.

Table 1. The email weaknesses and vulnerability related to the mailing process

CVEID	**VulnType**	**CVEID**	**VulnType**
CVE-2017-7440	CWE-1021	CVE-2006-1173	CWE-400
CVE-2001-1009	CWE-129	CVE-2006-4434	CWE-416
CVE-2003-0721	CWE-129	CVE-2001-0901	CWE-434
CVE-2004-0568	CWE-130	CVE-2002-0485	CWE-436
CVE-2000-0703	CWE-138	CVE-2002-0637	CWE-436
CVE-2001-0677	CWE-138	CVE-2002-1777	CWE-436
CVE-2003-0307	CWE-141	CVE-2005-0315	CWE-441
CVE-2000-0319	CWE-147	CVE-2001-0398	CWE-451
CVE-2001-0996	CWE-147	CVE-2002-1757	CWE-471
CVE-2000-0320	CWE-147	CVE-2002-0108	CWE-472
CVE-2003-1016	CWE-149	CVE-2005-1784	CWE-472
CVE-2000-0703	CWE-150	CVE-2005-1652	CWE-472
CVE-2002-0986	CWE-150	CVE-2005-1682	CWE-472
CVE-2002-0542	CWE-150	CVE-2000-1234	CWE-472
CVE-2004-0162	CWE-151	CVE-2008-3663	CWE-614
CVE-2003-1015	CWE-156	CVE-2006-2828	CWE-621
CVE-2002-0637	CWE-156	CVE-2007-0617	CWE-623
CVE-2005-2933	CWE-157	CVE-2005-4155	CWE-626
CVE-2005-4155	CWE-158	CVE-2006-3617	CWE-692
CVE-2002-1774	CWE-158	CVE-2001-1246	CWE-78
CVE-2002-1532	CWE-166	CVE-2008-5734	CWE-79
CVE-2004-2351	CWE-184	CVE-2003-1136	CWE-80
CVE-2005-1824	CWE-184	CVE-2002-1495	CWE-80
CVE-2005-4155	CWE-185	CVE-2003-1136	CWE-83
CVE-2002-1527	CWE-185	CVE-2005-0945	CWE-83

CVE-2008-1284	CWE-20	CVE-2004-1935	CWE-83
CVE-2006-5462	CWE-20	CVE-2002-1495	CWE-83
CVE-2002-1839	CWE-223	CVE-2005-2276	CWE-84
CVE-2009-4565	CWE-297	CVE-2005-0563	CWE-84
CVE-2007-5626	CWE-311	CVE-2002-0738	CWE-87
CVE-2001-1537	CWE-312	CVE-2001-1246	CWE-88
CVE-2001-1537	CWE-315	CVE-2006-2057	CWE-88
CVE-2007-5626	CWE-319	CVE-2002-0985	CWE-88
CVE-2002-0389	CWE-341	CVE-2006-2058	CWE-88
CVE-2001-0038	CWE-36	CVE-2004-0411	CWE-88
CVE-1999-1263	CWE-36	CVE-2004-0121	CWE-88
CVE-1999-1263	CWE-38	CVE-2006-2828	CWE-914
CVE-2001-0038	CWE-39	CVE-2002-1771	CWE-93

Table 1 shows email vulnerabilities (CVE entries) on the basis of CVE-Description property using the keyword "mail". The corresponding CWE entry is presented for each CVE entry (vulnerability). The number of the unrepeated CVE entries is (63) related to (48) CWE entries unrepeated. This query informs the security manager to consider the necessary measures to avoid them.

CASE 2: Obtain the email CVE entries on the basis of the ProductVendor keyword:

```
PREFIX :
<http://www.semanticweb.org/falak/ontologies/2018/0/20/untitled-
ontology-337#>
        SELECT DISTINCT *
        WHERE { ?CVEID  :CVEDescription ?CVEDescription;
                        :Component ?Component;
                        :ProductVendor ?ProductVendor;
                        :Version ?Version;
                        :RootCause ?RootCause;
                        :Impact ?Impact;
                        :Vector ?Vector;
                        :VulnType ?VulnType.
                ?VulnType :Name ?CWEName;
      filter( regex( ?ProductVendor, "mail")) } ORDER BY  ?VulnType
```

This query uses the ProductVendor property on the basis of the "mail" keyword, for the purpose of obtaining vulnerabilities CVEs with their related CWE entry(s) with properties Components, ProductVendor, Version, Root Cause, Impact, and Vector, and it declares the CWE entry properties ID, Name, and Description.

Table 2. The email vulnerabilities related to mailing process on the basis of the ProductVendor property

CVEID	VulnType	CVEID	VulnType
CVE-2001-1009	CWE-129	CVE-1999-1263	CWE-36
CVE-2000-0319	CWE-147	CVE-1999-1263	CWE-38
CVE-2005-2933	CWE-157	CVE-2006-1173	CWE-400
CVE-2002-1532	CWE-166	CVE-2006-4434	CWE-416
CVE-2005-1824	CWE-184	CVE-2001-0901	CWE-434
CVE-2002-1527	CWE-185	CVE-2005-0315	CWE-441
CVE-2008-1284	CWE-20	CVE-2001-0398	CWE-451
CVE-2009-4565	CWE-297	CVE-2008-3663	CWE-614
CVE-2001-1537	CWE-312	CVE-2002-1495	CWE-80
CVE-2001-1537	CWE-315	CVE-2002-1495	CWE-83

Table 2 shows email vulnerabilities (CVE entries) on the basis of the CVE-Description property using the "mail" keyword. The corresponding CWE entry is presented for each CVE entry (vulnerability). The result is 17 unrepeated CVE entries related to only 20 unrepeated CWE entries.

CASE 3. Testing email vulnerabilities on the basis of the Vector property:

```
PREFIX:<http://www.semanticweb.org/falak/ontologies/2018/0/20/untitled-
ontology-337#>
        SELECT DISTINCT *
        WHERE { ?CVEID  :CVEDescription ?CVEDescription;
                        :Component ?Component;
                        :ProductVendor ?ProductVendor;
                        :Version ?Version;
                        :RootCause ?RootCause;
                        :Impact ?Impact;
                        :Vector ?Vector;
                        :VulnType ?VulnType.
              ?VulnType :Name ?CWEName;
            filter( regex(?Vector, "mail")) } ORDER BY  ?VulnType
```

In this query the property Vector is used on the basis of the "mail" keyword, in order to present the email vulnerabilities CVE Entries with their related CWE entry(s) with properties Components, ProductVendor, Version, RootCause, Impact, and Vector, and it shows the CWE entry properties ID, Name, and Description.

Table 3. Mailing vulnerabilities on the basis of the Vector property

CVEID	VulnType	CVEID	VulnType
CVE-2017-7440	CWE-1021	CVE-2002-0108	CWE-472
CVE-2003-0721	CWE-129	CVE-2005-1784	CWE-472
CVE-2004-0568	CWE-130	CVE-2005-1652	CWE-472
CVE-2001-0677	CWE-138	CVE-2006-2828	CWE-621
CVE-2003-0307	CWE-141	CVE-2005-4155	CWE-626
CVE-2003-1016	CWE-149	CVE-2006-3617	CWE-692
CVE-2002-0542	CWE-150	CVE-2008-5734	CWE-79
CVE-2004-0162	CWE-151	CVE-2003-1136	CWE-80
CVE-2003-1015	CWE-156	CVE-2002-1495	CWE-80
CVE-2002-0637	CWE-156	CVE-2003-1136	CWE-83
CVE-2005-2933	CWE-157	CVE-2005-0945	CWE-83
CVE-2005-4155	CWE-158	CVE-2004-1935	CWE-83
CVE-2004-2351	CWE-184	CVE-2002-1495	CWE-83
CVE-2005-4155	CWE-185	CVE-2005-2276	CWE-84
CVE-1999-1263	CWE-36	CVE-2005-0563	CWE-84
CVE-1999-1263	CWE-38	CVE-2006-2057	CWE-88
CVE-2006-1173	CWE-400	CVE-2006-2058	CWE-88
CVE-2002-0485	CWE-436	CVE-2006-2828	CWE-914
CVE-2002-0637	CWE-436	CVE-2002-1771	CWE-93
CVE-2002-1757	CWE-471	CVE-2017-7440	CWE-1021

Table 3 shows email vulnerabilities (CVE entries) on the basis of the Vector property using the keyword "mail". The corresponding CWE entry is presented for each CVE entry (vulnerability). In this case, the number of unrepeated vulnerabilities is 32 that related to 29 unrepeated.

4. Analysis of the Results. The security manager can make correct decisions to avoid mailing vulnerabilities. He/she can fix the email vulnerabilities through affected factors, i. e., CVE Description, ProductVendor and Vector. A lot of vulnerabilities probably put the security system to various risks. He/she should exchange reports with the software developer and the security team to remove and avoid the vulnerabilities. Meanwhile, various kinds of audience should work as one team to minimize or entirely remove the risks. This high number of email vulnerabilities can put the security system and the networks to risk.

5. Conclusions. Ontologies provide comprehensive capabilities for constructing an interactive knowledge base through interpreting the concepts of definitions in the knowledge domain. The security manager can make correct decisions through analyzing the obtained knowledge concepts. The SPARQL queries support the security of the mailing process.

The approach presented here can be extended to suit different domain areas. The email systems domain has been chosen to eliminate the contradictions of domain area that exists in CWE and CVE databases.

As a result of this research, a knowledge base for email CWE entries has been developed. The use of this knowledge base has been briefly demonstrated for the improvement of the security of the mailing process.

REFERENCES

[1] Gupta B. B., N. A. G. Arachchilage, K. E. Psannis. Defending against phishing attacks: taxonomy of methods, current issues and future directions. *Telecommunication Systems*, **67** (2018), 247–267.

[2] Nahorney B. Email Threats 2017: An ISTR Special Report. Symantec, 2017. https://www.symantec.com/content/dam/symantec/docs/security-center/white-papers/istr-email-threats-2017-en.pdf, 30 January 2020.

[3] DuCharme B. Google's Learning SPARQL: Querying and Updating with SPARQL 1.1. O'Reilly Media, 2011.

[4] Common Vulnerabilities and Exposure. About CVE. https://cve.mitre.org/about/, 30 January 2019.

[5] MITRE: Solving Problems for a Safer World. https://www.mitre.org/, 25 November 2021.

[6] Congressional Task Force on Election Security: Final Report. January 2018. https://homeland.house.gov/imo/media/doc/TFESReport.pdf, 25 November 2021.

[7] Symantec. ISTR: Internet Security Threat Report, Volume **24** (2019). https://docs.broadcom.com/doc/internet-security-threat-report-volume-24-en.

[8] Hammar K. Content Ontology Design Patterns: Qualities, Methods, and Tools. Linköping University, 2017.

[9] Kaspersky Lab. The Great Bank Robbery. Version 2.1, February 2015.

[10] Horrige M., S. Brandt. A Practical Guide to Building OWL Ontologies Using Protégé 4 and CO-ODE Tools: Edition 1.3. The University of Manchester, 2011. http://mowl-power.cs.man.ac.uk/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf, 25 November 2021.

[11] OWL Web Ontology Language Reference, W3C Recommendation. https://www.w3.org/TR/owl-ref/, 30 January 2020.

[12] Slimani T. Ontology Development: A Comparing Study on Tools, Languages and Formalisms. *Indian Journal of Science and Technology*, **8** (2015), No 24, 1–12.

[13] Sheng S., M. Holbrook, P. Kumaraguru, L. F. Cranor, J. Downs. Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions. In: Proceedings of CHI'10, SIGCHI Conference on Human Factors in Computing Systems, 2010, 373–382.

[14] CWE: Common Weakness Enumeration. https://cwe.mitre.org/about/faq.html, 30 January 2020.

Falak Ussien Hasan
Faculty of Mathematics and Informatics
St. Kliment Ohridski University of Sofia
5, James Bourchier Blvd
1164 Sofia, Bulgaria
e-mail: falakhasan2014@yahoo.com

Received November 14, 2019
Final Accepted June 7, 2020