Numerical and Statistical Analysis on Information Security Using R: Interoperability Perspective

Kerim Goztepe

Sakarya University, Institute of Science, Dept.of Industrial Engineering, 54330, Sakarya-Turkey

E-mail: kerimgoztepe@gmail.com

ABSTRACT

Over the past two decades, information and communication technology (ICT) have gained enormous effect on social-business life more than ever. ICT has become increasingly important not only for life but also national security. Interoperability among ICTs is also a vital topic to support national security. The term “interoperability” covers a nation’s information telecommunication networks, critical infrastructures, intelligence and military systems which are all vulnerable to growing number of cyber-attacks from inside and outside. The purpose of this study is to examine and statistically analyze the importance of interoperability applying cost, cyber espionage (CE), cyber warfare (CW), hacktivism (HA) and cyber-crime (CC) criteria based on available information about cyber-attacks to governmental organizations or critical infrastructure. For this goal, obtained dataset from opensource has been analyzed. Firstly, Shapiro-Wilk and Anderson-Darling is applied for checking normality of the data. After this, a regression model is developed to see the relation between cost and cyber-attacks types. The results have shown that cyber-attacks affect interoperability environment and cause losing great financial resources. Thus, not only governments but also business world should pay more attention to develop advanced security integration models.

KEYWORDS

Information and communication technology (ICT), Cyber-attacks, Statistics, Interoperability

1 INTRODUCTION

Today many information security experts approve information security (or cyber security) as a matter of global interest and importance for organizations, firms and governments. According to Klimburg [1], more than fifty states have officially published cyber security strategy document outlining their official position to cyberspace events.

The main purpose of information security is to assure business continuity and minimize possible damage by determining some technical issues for information systems [2]. Another researcher, Whitman (and Mattord), characterize information security as the process of information protect and vital information systems with the hardware [3]. Basically, ICT security handles with the protection of the whole technology-based systems on which data is commonly stored and/or transmitted in an organization or firm. Information security is sometimes used instead of cyber security because many cyber security researchers use it interchangeably with the term information security [4].

Availability, integrity, and confidentiality are key definitions for cyber security. Information experts usually explain the vulnerability of cyber systems against to viruses, denial-of-service attacks, worms, malware and errors made by users [5], [6]. According to the latest reports; total number, type and sophistication cyber security threats level to information systems are boosting. Motivations of these attacks are varied, but attackers deploy malicious activities even if they do not think the information held on the target systems is valuable [7], [8].

Recent studies have indicated that the number of cyber attackers and related to these total attacks are being increased [9]. This circumstance forces organizations and states to take more comprehensive security measures in terms of interoperability. It is believed, the globalization and gaining complexity of today’s networks have made it nearly unable to cope with cyber threats for any organization [10], [11].
Interoperability is a term that defines the ability of a computer system to run different software from varied vendors and to collaborate with other information systems across local or wide-area networks even physical architecture and operating systems are not similar [12]. In another definition, interoperability policy defines the term as the ability for governments or organizations (NATO, EU etc.) to work together efficiently and effectively to obtain determined targets [13].

In order to control cyber-attacks for a better and manageable interoperability, governmental organizations should collaborate with each other. This proposed study presents a short statistical analysis on cyber security that cyber-attacks are a crucial issue for cyber defense.

2 INCREASING THREATS FOR ICT SYSTEMS: CYBER-ATTACKS

It is claimed that cyber-attacks are the greatest new threats in today’s information world. This truth is a reality for not only the great break that cyber-attacks can affect whole society but also the number of different issues that is involved in this world [14]. This is obviously a modern global threat for any state and a policy should be developed for it. For instance, "Aurora" code-named attacks in 2010; the "Conficker" called worms; Stuxnet worm aiming a nuclear program, and many new cyber-attack methods are all new kind of threats for the global world [15], [16]. Organizations must take precautions to cyber-threats, inaction cannot be possible.

In this paper, four main cyber-attack types: cyber espionage (CE), cyber warfare (CW), hacktivism (HA) and cyber-crime (CC) have been evaluated using best-known data science software R. These are explained briefly in the following sentences:

**Cyber espionage:** Cyber espionage expresses the stealing secret information that digitally formatted from computers and IT networks. For instance, “Analysis of the Gauss” has designed to collect information and sends to its owners [17].

**Cyber warfare:** Cyber warfare defines internet-based attacks to networks and information systems. Its main aim is to paralyze official websites and networks, disable social services, banking, transportation or irrigation systems [18].

**Hacktivism:** Hacktivism is the act of hacking a website or computer network. Hacktivist tries to declare a political or social message applying hactivism [19].

**Cybercrime:** It is a crime type executed by computers. This crime may be done by different cyber-attack tools as hacking, phishing, spamming, etc. [19].

3 INTEROPERABILITY

The latest conflicts around the world exposed the need for compatible equipment, standard defense terminology and well-established combat procedures to ensure interoperability between the services in the military [20]. Most of the defense experts have same idea that new defense technologies have caused in increasingly dynamic, unpredictable and complex operations in this new era. In this environment, military decision makers from different forces have to filter and analyze information from multiple sources [21], [22]. Beside decision-making, problem-solving and sense-making are more composite and more crucial in today's' battlefield [20]. Know-how and expertise are also significant issues in a military organization’s ability to achieve knowledge management in a combat environment. Thus, interoperability appeared as a fundamental principle of the joint operation of army, navy, and airforce [23]. Australian Defense Force authorities define interoperability as: “established system capabilities, units or forces to supply and accept services from different arrangements, divisions or forces and to use the services so exchanged to enable them to operate capably” [24]. In this perspective, any of military units should have the capability to operate at every point of the world with the ability of “interoperability for joint and coalition operations”.

**Electromagnetic Spectrum Operations:** Net-centric warfare is a part of interoperability and uses electromagnetic spectrum [25], [26].
Electromagnetic spectrum operations (EMSO) allow necessary resource for the implementation of the wireless technology (Fig.1). From this context, understanding the operational process in planning, managing, and executing is vital for all combat functions [27]. Today’s cyber-attacks aim to obtain secret military information, to damage command control systems or to take control of some combat system secretly. Operation commander should be aware of possible dangers for EMSO during operation planning.

4 STATISTICAL ANALYSIS on CYBER-attacks in TERMS of INTEROPERABILITY

Computers play a significant role in not only for business but also for interoperability of governmental organizations [28], [29]. They are responsible for the communication, control of critical infrastructure and most of the military technological systems. It is a known fact that cyber-attacks effect the stability of interoperability in today's world. In this study, we applied a statistical analysis in order to see effects of cyber-attacks and to model possible cost of it using regression analysis in R.

**R Software:** R is a data analysis software usually applied for statistical applications and obtaining detailed graphics of data which is created by Ross Ihaka and Robert Gentleman [30]. R is not only a statistical software but also conceived as a derivative language of S [31]. R has GNU General Public License and distributed freely. R development core team (consisted of several statisticians) is responsible for development and distribution of R. Researchers can find different R forms that run under Linux, Unix, Windows and Mac [32].

We evaluated four cyber-attack types related with interoperability. These are code execution (CEX), Denial of Service (DoS), gain privilege (GP) and cross-site scripting (XSS) (Table 1).

**Table 1:** Evaluated cyber-attack types

<table>
<thead>
<tr>
<th>Cyber-attack type</th>
<th>Evaluated for</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code execution (CEX)</td>
<td>Cyber espionage</td>
</tr>
<tr>
<td>DoS (Denial of Service)</td>
<td>Cyber warfare</td>
</tr>
<tr>
<td>Gain privilege (GP)</td>
<td>Cyber-crime</td>
</tr>
<tr>
<td>XSS (Cross-site Scripting)</td>
<td>Hacktivism</td>
</tr>
</tbody>
</table>

The data are derived from open sources as shown in Fig.2 and Appendix B.

**Figure 2.** Evaluated cyber-attack types that affect interoperability of governmental organizations (Source: https://www.cvedetails.com/vulnerabilities-by-types.php)

As can be seen Fig. 2., total DoS attacks between years 1999-2017 is 19078, code execution is 25667, gain privilege is 4356 and XSS are 10718. Other detailed information about information security threats can be seen in
Appendix B which is real dataset obtained from opensource.

4.1 STATISTICAL ANALYSIS of DATA USING R

Normality test: We applied normality test for the criteria CEX, DoS, GP, and XSS. One of the most known normality test, The Shapiro-Wilk, is used here for calculation procedure [33]. The Shapiro-Wilk calculates a W statistic value that tests if a given sample \( x_1, x_2, ..., x_n \) comes from a normal distribution.

\[
W = \frac{\left( \sum_{i=1}^{n} a_i x_{(i)} \right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}
\]

In Eq.1, \( x_{(i)} \) are the ordered sample values and \( a_i \) are the constants.

R code for calculation, results, and data plot (see Fig.3) are given below.

```r
# Cyber attacks data transferred to R environment as cyber. Appendix B data is used here.
> library(readxl)
> cyber<- read_excel("file location")
> View(cyber)
# Change format of our data cyber as data.frame to cyberdf
> cyberdf<-data.frame(cyber)
> attach(cyberdf)
> head(cyberdf)
# Now we can make normality test for DoS data.
> shapiro.test(DoS)
Shapiro-Wilk normality test
data:  DoS
W = 0.96971, p-value = 0.7707
> qnorm(DoS); qqline(DoS)
```

![Normal Q-Q Plot](image)

**Figure 3.** Normality plot of DoS

Basic statistical data (measure of central tendency) of DoS criteria is given below [summary(DoS)]. Minimum value of data is 177 DoS attacks that refer to the year 1999. Maximum attacks have been done in 2016.

```r
> summary(DoS)
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
177.0  538.5  1035.0  1004.1  1414.5  2029.0
```

Over the years, there is an absolute increase in DoS attacks. The relationship between year and DoS attacks figure (Fig.4) and R code is given below.

```r
> library(ggplot2)
> p<-ggplot(cyberdf, aes(x=DoS, y=Year))
+ geom_point(shape=23, color="red", size=4)+geom_smooth(method=lm, linetype="dashed",color="darkred", fill="blue")
> p+theme_bw()
> p
```

![Year-DoS relationship](image)

**Figure 4.** Year-DoS relationship

Normality test for CEX is given below. W value is 0.9529 and p-value is 0.4425 according to test result.

```r
> shapiro.test(Code.Execution)
Shapiro-Wilk normality test
data:  Code.Execution
W = 0.95293, p-value = 0.4425
```

Histogram is another way of getting information about data. A histogram is a plot that lets researchers discover, and present, the underlying data frequency distribution. By this way, someone can make a detailed (outliers, skewness, etc) inspection of the data. Histogram of CEX is shown below (Fig.5)

```r
> hist(Code.Execution)
```

![Histogram](image)

**Figure 5.** Histogram of CEX
When we examine Fig. 6, it can be seen that there is a decrease for CEX attacks in recent years. It is possible to conclude from summary data above that maximum value is not belongs to latest years.

Data summary and normality tests of GP and XSS is given below.

```
> summary(Gain.Privileges)
  Min. 1st Qu. Median  Mean 3rd Qu. Max.
   103  184.5  220.0  229.3  240.5 601.0

> summary(XSS)
  Min. 1st Qu. Median  Mean 3rd Qu. Max.
    2.0  245.5  605.0  564.1  796.5 1302.0

> shapiro.test(Gain.Privileges)
Shapiro-Wilk normality test
data:  Gain.Privileges
W = 0.74747, p-value = 0.0002168

> shapiro.test(XSS)
Shapiro-Wilk normality test
data:  XSS
W = 0.95298, p-value = 0.4434
```

Normality test performs a hypothesis test to understand whether or not the observations follow a normal distribution. In this test, the hypotheses are:

- **H₀**: data follow is a normal distribution.
- **H₁**: data do not follow a normal distribution.

When we assume the significance level at 0.05 than we cannot reject the null hypothesis about the normal distribution for DoS (p-value=0.7707), CEX (p-value=0.4425) and XSS (p-value=0.4434). Because p value is greater than 0.05.

For GP criteria, we can conclude that GP data is not normally distributed, because p-value=0.0002168<0.05. We reject H₀ and accept H₁, data do not follow a normal distribution. It is possible to examine result graphically in Fig. 7.

**Anderson-Darling test:** It is possible to deploy different tests in R environment in order to check normality. Anderson-Darling (A-D) test is one of the normality tests for testing normality of data. The data from the A-D test gives us two results: A and p-value. A value represents the result of Anderson-Darling test statistic [34].

The Anderson-Darling test is defined as:

- **H₀**: The data follow a specified distribution.
- **H₁**: The data do not follow the specified distribution.

The Anderson-Darling test statistic is defined as

$$A^2 = -N - S$$

where
\[ s = \sum_{i=1}^{N} \frac{(2i-1)}{N} [\ln(F(Y_i)) + \ln(1 - F(Y_{N+i}))] \]  

(2)

Shapiro-Wilk test results inform us GP criteria is not normally distributed. We deployed Anderson-Darling (A-D) test to check normality of GP.

A-D test result of GP criteria is given below. While p-value(0.0004037)<0.05 we reject \( H_0 \).

\[
\text{Shapiro-Wilk test results for GP criteria}
\]

Regression analysis: A regression analysis has been made in order to explain the simulated cost of cyber-attacks types (CEX, DoS, GP, XSS) in this study. Our aim is to indicate that cyber security is a crucial issue and governmental organizations may lose billions of dollars in cyber incidents. Besides, this may cause a dilemma in terms of interoperability. Regression is a multiple linear model and explains the linear relationship between the response variable (Y) and k (>1) multiple regressors:

\[
x_j's \quad Y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + \epsilon
\]

(3)

where regression coefficients \( \beta_0, \beta_1, ..., \beta_k \) are. The coefficient \( \beta_0 \) explains the intercept of the model and the coefficients \( \beta_j's, j = 1 \) to \( k \) shows the slopes of the regressor. Each slope coefficient \( \beta_j \) measures the expected change in \( Y \) per unit change in \( x_j \), when the other regressors are held constant [35], [36].

A regression analysis has been done to explain the cost - cyber-attacks relationship. There is no exact data related with cyber-attacks in literature. Data taken from Ponemon Institute [37] is simulated under R and used for regression model. The following R code has been executed in order to simulate cost of cyber-attacks.

\[
\text{R code for simulating cost of cyber-attacks}
\]

After simulation of “cost” values, regression model “cybermodel”, is created under R.

\[
\text{cybermodel <- lm(Cost ~ Code.Execution + DoS + Gain.Privileges + XSS, data=cyberdf)}
\]

According to Eq.3 the relationship between \( Y \) (cost) and criteria CEX, DoS, GP, XSS is given below.

\[
\text{Cost} = 5897807.68 + 21.93 \times \text{CEX} + 2742.25 \times \text{DoS} + -3520.15 \times \text{GP} + -424.43 \times \text{XSS}
\]

(4)

## Researcher can see model, writing its name in R console.

\[
\text{cybermodel}
\]

(5)

\[
\text{Call: lm(formula = Cost ~ Code.Execution + DoS + Gain.Privileges + XSS, data = cyberdf)}
\]

Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>Code.Execution</th>
<th>DoS</th>
<th>Gain.Privileges</th>
<th>XSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>5987807.68</td>
<td>21.93</td>
<td>2742.25</td>
<td>-3520.15</td>
</tr>
<tr>
<td>Gain.Privileges</td>
<td>-424.43</td>
<td>21.93</td>
<td>2742.25</td>
<td>-3520.15</td>
</tr>
</tbody>
</table>
We use the R-squared statistic ($R^2$) to measure how well the model is fitting the data applied in the model [35]. It is a measure of the linear relationship between our predictor variable and our response. It always gets a number between 0 and 1. If the result is near 0, it means that model does not explain the variance in the response variable well and if the result is close to 1 then it explains the observed variance in the response variable best.

In our proposed model, the result of $R^2$ is 0.8948. It means 89% of the variance found in the response variable (cost) can be explained by the predictor variables (CEx, DoS, GP, XSS). $R^2$ always increase as more variables are added to the model.

Adjusted $R^2$ is the preferred measure how well model fit a curve or line. For a proposed regression model; if a researcher adds more useless variables, Adjusted $R^2$ will decrease. The numerical value of Adjusted $R^2$ will increase by adding more useful variables only. Another important point, Adjusted $R^2$ will always be less than or equal to $R^2$ [35].

Graphic output of model: R has great ability to produce proposed model graphics natively. Residuals vs Fitted plot checks non-linearity of model [38]. If data has equally spread residuals around a horizontal line without distinct patterns, that means data has linear relationship. In our model data has linear relationship. Normal Q-Q plot depictions if residuals are normally distributed. We can say that our data mostly normal distributed. Scale-Location plot shows residuals’ spread along the predictors. We see a horizontal line with equally spread points in Fig.9/Scale-Location which proves model health. Residuals vs Leverage plot assists us to observe influential cases in model.

### Interpretation of model

**p-value:** As can be seen in above results, the relationship between the variables in the model is statistically significant (p=1.036e-06<0.01).

**Residual Standard Error(RSE):** RSE represents the quality of a linear regression fit [35]. In theory, every linear model is assumed to contain an error “E” term. The RSE is the average amount that the response will deviate from the true regression line. In our model, the yearly actual cyber-attack cost can deviate from the true regression line by approximately 427800 USA dollars.

**Multiple R-squared, Adjusted R-squared:** We use the R-squared statistic ($R^2$) to measure how well the model is fitting the data applied in the model [35]. It is a measure of the linear relationship between our predictor variable and our response. It always gets a number between 0 and 1. If the result is near 0, it means that model
between our variables [35]. If the F-statistic value is far from 1, then model quality is better. In our proposed model, the F-statistic is 29.77 which is relatively larger than 1.

The relationship between cyber-attacks costs and types of cyber-attacks is depicted in Fig.10. If the model fits the data well, this equation can be used to predict cost for specific values of the X variables, or find the settings for the X variables that correspond to a desired value or range of values for cost. It can be seen that cost increases as number of attacks change for every criteria (cost is given in USA dollars).

```r
## R code of Fig.10
> library(ggplot2)
> costvscyber <- ggplot(cyberdf, aes(x=Cost)) +
  geom_line(aes(y=Code.Execution, color="Code.Execution")) +
  geom_line(aes(y=DoS, color="DoS")) +
  geom_line(aes(y=Gain.Privileges, color="Gain.Privileges")) +
  geom_line(aes(y=XSS, color="XSS"))
> costvscyber+theme(text=element_text(size=16))
```

Figure 10. Relationship between cost and types of cyber attacks.

Other useful output about model is given in Appendix A with R code (model coefficients, predicted values, anova table and information about “covariance matrix for model parameters”)

5 CONCLUSION

As mentioned many experts, the internet and digital technologies are converting people, states, and business by driving economic growth, connecting world and providing new ways to communication. Rising of internet affected not only millions of people but also governments, social life, organizations and armies. However, a new and important topic has appeared; cyber security.

This paper presents a short numerical and statistical analysis on cyber-attacks that affects interoperability. The aim of this study is to make curious cyber decision makers (both military and civil) about main cyber-attack effects on interoperability. For this purpose, a real and simulated data (for cost criteria) has been used in R for analyzing. A regression model has been developed in order to see the interaction between the cost and some well-known cyber-attacks. According to the dataset, increasing cyber-attack trend can be seen in 3-D criteria relationship (Fig.11, Fig.12)

```r
##R code of Fig.11
library(rgl)
x <- Cost
y <- Code.Execution
z <- DoS
plot3d(x=x, y=y, z=z, size = 4, type = "s", col = "blue")
```

Figure 11. 3-D interact between Cost, CE and DoS
The internet has increased number of computers and infrastructures including complex interactions among cyber components. Cyber-attacks affect interoperability environment and this causes to lose great financial resources. Governments have to develop advanced security integration models. It is hoped that this study might trigger national needs for a better cyber security complexities involved in national operating system for interoperability environment.

REFERENCES


APPENDIX A

Other useful information about model

```r
> coefficients(cyber) # model coefficients
5897807.68452  21.93454     2742.25124 -3520.14668 -424.42941

> confint(cyber, level=0.95) # CIs for model parameters
     2.5 %       97.5 %
(Intercept)     5331190.4845 6464424.8845
Code.Execution  -700.4295     744.2986
DoS                1799.5813    3684.9212
Gain.Privileges  -7440.1432     399.8499
XSS              -2180.4916    1331.6328

> fitted(cyber) # predicted values
            1         2         3         4         5         6         7         8         9        10        11
6022219 6116131 6220169 6487441 6391394 6906573 7119926 7205964 7744689 7395747 7637783
12        13        14        15        16        17        18        19
7866310 8351998 8635720 8670650 9001376 9238231 9168905 8874203

> anova(cyber) # anova table

Analysis of Variance Table

Response: Cost
  Df  Sum Sq Mean Sq F value Pr(>F)
Code.Execution  1 5.6854e+12 5.6854e+12 31.0676 6.858e-05 ***
DoS            1 1.5353e+13 1.5353e+13 83.8962 2.740e-07 ***
Gain.Privileges 1 7.0221e+11 7.0221e+11  3.8372   0.07035 .
XSS            1 4.9176e+10 4.9176e+10  0.2687   0.61229
Residuals      14 2.5620e+12 1.8300e+11

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> vcov(cyber) # covariance matrix for model parameters
   (Intercept)  69792908373 -10002504.34 15670166.56 -223567111.2 19716987.3
Code.Execution -10002504  113434.20   48915.34  -239065.2  -243929.9
DoS            15670167  48915.34   193175.09   -675246.6 -214382.6
Gain.Privileges -223567111 -239065.21  -675246.59  3340436.1  813367.0
XSS            -19716987  -243929.95  -214382.59   813367.0  670365.4
```
# APPENDIX B

Information security report

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>894</td>
<td>177</td>
<td>208</td>
<td>172</td>
<td>112</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>7</td>
<td>25</td>
<td>16</td>
<td>103</td>
<td>2</td>
<td>5444675</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>1020</td>
<td>257</td>
<td>206</td>
<td>208</td>
<td>208</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>20</td>
<td>48</td>
<td>19</td>
<td>139</td>
<td>2</td>
<td>5782803</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>1677</td>
<td>402</td>
<td>403</td>
<td>297</td>
<td>403</td>
<td>7</td>
<td>34</td>
<td>123</td>
<td>83</td>
<td>36</td>
<td>220</td>
<td>2</td>
<td>6244406</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>2156</td>
<td>497</td>
<td>553</td>
<td>435</td>
<td>553</td>
<td>2</td>
<td>41</td>
<td>200</td>
<td>103</td>
<td>127</td>
<td>74</td>
<td>199</td>
<td>2</td>
<td>6785425</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>1526</td>
<td>381</td>
<td>476</td>
<td>370</td>
<td>476</td>
<td>2</td>
<td>49</td>
<td>129</td>
<td>60</td>
<td>61</td>
<td>69</td>
<td>144</td>
<td>2</td>
<td>6965342</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>2450</td>
<td>580</td>
<td>614</td>
<td>410</td>
<td>614</td>
<td>3</td>
<td>148</td>
<td>291</td>
<td>110</td>
<td>12</td>
<td>145</td>
<td>96</td>
<td>134</td>
<td>5</td>
<td>7032664</td>
<td>13</td>
</tr>
<tr>
<td>2005</td>
<td>4935</td>
<td>838</td>
<td>1626</td>
<td>657</td>
<td>1626</td>
<td>21</td>
<td>604</td>
<td>786</td>
<td>202</td>
<td>15</td>
<td>289</td>
<td>261</td>
<td>221</td>
<td>11</td>
<td>7306158</td>
<td>25</td>
</tr>
<tr>
<td>2006</td>
<td>6610</td>
<td>893</td>
<td>2719</td>
<td>662</td>
<td>2719</td>
<td>91</td>
<td>967</td>
<td>1302</td>
<td>322</td>
<td>8</td>
<td>267</td>
<td>267</td>
<td>184</td>
<td>18</td>
<td>7313047</td>
<td>97</td>
</tr>
<tr>
<td>2007</td>
<td>6520</td>
<td>1100</td>
<td>2601</td>
<td>953</td>
<td>2601</td>
<td>95</td>
<td>706</td>
<td>883</td>
<td>339</td>
<td>14</td>
<td>267</td>
<td>322</td>
<td>242</td>
<td>69</td>
<td>7369950</td>
<td>1251</td>
</tr>
<tr>
<td>2008</td>
<td>5632</td>
<td>894</td>
<td>2310</td>
<td>699</td>
<td>2310</td>
<td>128</td>
<td>1101</td>
<td>807</td>
<td>363</td>
<td>7</td>
<td>288</td>
<td>270</td>
<td>188</td>
<td>83</td>
<td>7573773</td>
<td>1936</td>
</tr>
<tr>
<td>2009</td>
<td>5736</td>
<td>1035</td>
<td>2185</td>
<td>700</td>
<td>2185</td>
<td>188</td>
<td>963</td>
<td>851</td>
<td>322</td>
<td>9</td>
<td>337</td>
<td>302</td>
<td>223</td>
<td>115</td>
<td>7663233</td>
<td>2108</td>
</tr>
<tr>
<td>2010</td>
<td>4651</td>
<td>1102</td>
<td>1714</td>
<td>678</td>
<td>1714</td>
<td>342</td>
<td>520</td>
<td>605</td>
<td>275</td>
<td>8</td>
<td>234</td>
<td>282</td>
<td>237</td>
<td>86</td>
<td>7904878</td>
<td>1469</td>
</tr>
<tr>
<td>2011</td>
<td>4155</td>
<td>1221</td>
<td>1334</td>
<td>770</td>
<td>1334</td>
<td>351</td>
<td>294</td>
<td>467</td>
<td>108</td>
<td>7</td>
<td>197</td>
<td>409</td>
<td>206</td>
<td>58</td>
<td>7957099</td>
<td>564</td>
</tr>
<tr>
<td>2012</td>
<td>5297</td>
<td>1425</td>
<td>1457</td>
<td>842</td>
<td>1457</td>
<td>423</td>
<td>242</td>
<td>758</td>
<td>122</td>
<td>13</td>
<td>343</td>
<td>389</td>
<td>250</td>
<td>166</td>
<td>8252564</td>
<td>615</td>
</tr>
<tr>
<td>2013</td>
<td>5191</td>
<td>1454</td>
<td>1186</td>
<td>859</td>
<td>1186</td>
<td>366</td>
<td>156</td>
<td>650</td>
<td>110</td>
<td>7</td>
<td>352</td>
<td>511</td>
<td>274</td>
<td>123</td>
<td>8388240</td>
<td>205</td>
</tr>
<tr>
<td>2014</td>
<td>7946</td>
<td>1597</td>
<td>1573</td>
<td>850</td>
<td>1573</td>
<td>420</td>
<td>304</td>
<td>1105</td>
<td>204</td>
<td>12</td>
<td>457</td>
<td>2104</td>
<td>239</td>
<td>264</td>
<td>8908746</td>
<td>399</td>
</tr>
<tr>
<td>2015</td>
<td>6452</td>
<td>1792</td>
<td>1820</td>
<td>1078</td>
<td>1820</td>
<td>749</td>
<td>213</td>
<td>758</td>
<td>146</td>
<td>12</td>
<td>577</td>
<td>748</td>
<td>367</td>
<td>245</td>
<td>8933563</td>
<td>118</td>
</tr>
<tr>
<td>2016</td>
<td>6435</td>
<td>2029</td>
<td>1494</td>
<td>1326</td>
<td>1494</td>
<td>717</td>
<td>94</td>
<td>495</td>
<td>98</td>
<td>15</td>
<td>444</td>
<td>840</td>
<td>601</td>
<td>86</td>
<td>9390760</td>
<td>1</td>
</tr>
<tr>
<td>2017</td>
<td>5989</td>
<td>1404</td>
<td>1292</td>
<td>987</td>
<td>1292</td>
<td>319</td>
<td>116</td>
<td>591</td>
<td>112</td>
<td>5</td>
<td>272</td>
<td>685</td>
<td>185</td>
<td>141</td>
<td>9838109</td>
<td>3</td>
</tr>
</tbody>
</table>

XSS: Cross-site Scripting  
CSRF: Cross-Site Request Forgery