Cyber Crimes in India

Project Report

Hitika

Contents

Chapter 1

Analysis of Cyber crimes under IT Act and IPC Act

1.1 Introduction

1.1.1 Information Technology ACT, 2000

The IT Act was introduced in 17 October 2000. It is the primary law in India dealing with cyber crime and electronic commerce. The original Act contained 94 sections, divided into 13 chapters and 4 schedules. The laws apply to the whole of India. If a crime involves a computer or network located in India, persons of other nationalities can also be indicted under the law. The sections of IT Act included in the analysis are

- Tampering Computer Source Documents (Section 65)
- Computer Related Offenses (Sec 66 & Sec 66 B to E)(TOTAL)
- Cyber Terrorism (Section 66F)
- Publication/ transmission of obscene / sexually explicit act in electronic form

The figure 3.1 shows the rate of cases of Information Technology Act, 2000 being registered in INDIA per 1 lakh population in years 2016-21. One must note that the Computer Related Offenses (Sec 66 & Sec 66 B to E) are reported mostly.

1.1.1.1 Computer Related Offenses (Sec 66 & Sec 66 B to E)

This includes punishment regarding sending offensive messages through communication service,Punishment for dishonestly receiving stolen computer resource or communication, Punishment for identity theft, Punishment for cheating by personation by using computer resource, Punishment for violation of privacy, Penalty and compensation for damage to computer, computer system, etc.

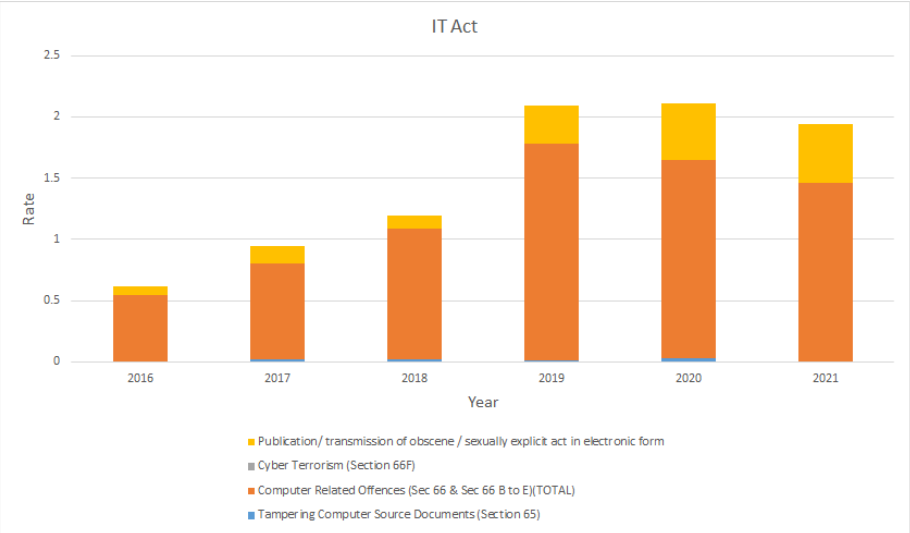

Figure 1.1: rates section-wise yearly it-act cases

1.1.2 Indian Penal Code , 1862

The Indian Penal Code (IPC) is the official criminal code of India. It is a comprehensive code intended to cover all substantive aspects of criminal law. It came into force in India during the British rule in 1862. The sections of IPC included in the analysis are

- Data theft (Sec.379 to 381)
- Fraud (Sec.420 r/w Sec.465,468-471 IPC)
- Cheating (Sec.420)
- Forgery (Sec.465, 468 & 471)

The figure 3.2 shows the rate of cases of Indian Penal Code (here only those sections are included which include punishments to cyber crimes) being registered in INDIA per 1 lakh population in years 2016-21. One must note that the FRAUD cases are reported mostly.

1.1.2.1 Fraud (Sec.420 r/w Sec.465,468-471 IPC)

This include that whoever cheats and thereby dishonestly induces the person deceived to deliver any property to any person, or to make, alter or destroy the whole or any part of a valuable security, or anything which is signed or sealed, and which is capable of being converted into a valuable security, shall be punished with imprisonment and shall also be liable to fine.

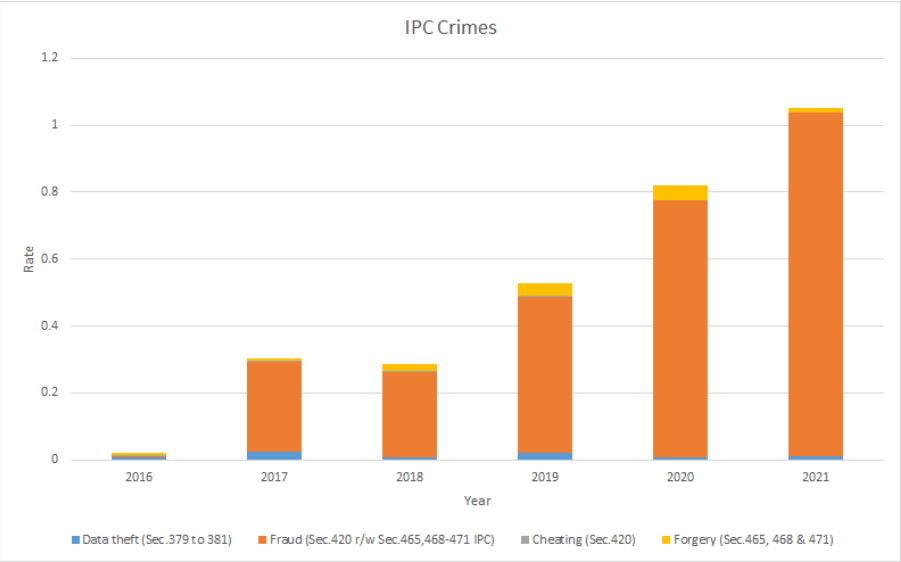

Figure 1.2: rates section-wise yearly ipc-act cases

1.2 Percentage share of major crimes under IT Act and IPC

```
##      year       CRO     Total per_share
## [1,] 2016 0.5351648 0.9667975   55.3544
## [2,] 2017 0.7844598 1.6915400   46.3755
## [3,] 2018 1.0685517 2.0589702   51.8974
## [4,] 2019 1.7652248 3.3300190   53.0095
## [5,] 2020 1.6198963 3.6969307   43.8173
## [6,] 2021 1.4566373 3.8746626   37.5939
```

```
##      year       Fraud     Total per_share
## [1,] 2016 0.004395604 0.9667975    0.4547
## [2,] 2017 0.268988693 1.6915400   15.9020
## [3,] 2018 0.253366380 2.0589702   12.3055
## [4,] 2019 0.465976884 3.3300190   13.9932
## [5,] 2020 0.768054263 3.6969307   20.7755
## [6,] 2021 1.024510127 3.8746626   26.4413
```

The above data table shows that Computer Related Offenses (Sec 66 & Sec 66 B to E) ,(here , Computer Related Offenses (Sec 66 & Sec 66 B to E) is named CRO) , has been contributing about 48% and Fraud contribute about 15% of the total cyber crimes each year as per the NCRB data for cyber crimes from year 2016 - 2021. Hence Computer Related Offenses (Sec 66 & Sec 66 B to E) is the major contributor to the cyber crimes in INDIA.

1.3 Randomness and Association

1.3.1 1-sample Run test for randomness

1-sample run test is used to test whether the observations are random

1.3.1.1 Computer related offenses (Sec 66 & Sec 66 B to E)

```
##
##  Runs Test for Randomness
##
## data:  cro_c
## runs = 2, m = 3, n = 3, p-value = 0.2
## alternative hypothesis: true number of runs is not equal the expected number
## sample estimates:
## median(x)
##     17028
```

This shows that year wise incidence of cases of Computer related offenses are random .

1.3.1.2 Fraud

```
##
##  Runs Test for Randomness
##
## data:  fr_c
## runs = 2, m = 3, n = 3, p-value = 0.2
## alternative hypothesis: true number of runs is not equal the expected number
## sample estimates:
## median(x)
##    4849.5
```

This shows that year wise incidence of cases of Fraud are random .

1.3.2 Kolmogorov Smirnov test

This test is used for testing whether both the samples come from same distribution

Test statistics,

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)|,$$

where $F_{1,n}$ and $F_{2,m}$ are the empirical distribution functions of the first and the second sample respectively, and sup is the supremum function.

```
##
##  Exact two-sample Kolmogorov-Smirnov test
##
## data:  y2016 and y2021
## D = 0.14286, p-value = 0.9333
## alternative hypothesis: two-sided
```

```
##
##  Exact two-sample Kolmogorov-Smirnov test
##
## data:  y2017 and y2021
## D = 0.14286, p-value = 0.9366
## alternative hypothesis: two-sided
```

```
##
##  Exact two-sample Kolmogorov-Smirnov test
##
## data:  y2018 and y2021
## D = 0.10714, p-value = 0.9956
## alternative hypothesis: two-sided
```

```
##
##  Exact two-sample Kolmogorov-Smirnov test
##
## data:  y2019 and y2021
## D = 0.14286, p-value = 0.9301
## alternative hypothesis: two-sided
```

```
##
##  Exact two-sample Kolmogorov-Smirnov test
##
## data:  y2020 and y2021
## D = 0.14286, p-value = 0.9361
## alternative hypothesis: two-sided
```

This shows that the distribution of computer related offenses throughout the states remains the same in the years 2016 - 2021 i.e. cases are concentrated in specific states and are distrubuted similarily each year.

1.3.3 Kendall Tau Measure of Association

It measures the relationship between two columns of ranked data.

Null hypothesis, The columns are not associated with each other against, The columns are associated with each other

Test statistics,

$$\tau = \frac{(\text{number of concordant pairs}) - (\text{number of discordant pairs})}{(\text{number of pairs})}$$

Test is to reject null hypothesis for

```
##
##  Kendall's rank correlation tau
##
## data:  cro_c and fr_c
## T = 11, p-value = 0.2722
## alternative hypothesis: true tau is not equal to 0
```

```
## sample estimates:
##       tau
## 0.4666667
```

This shows that there is no association between cyber crimes under Computer Related Offenses (Sec 66 & Sec 66 B to E) and Fraud.

1.4 Testing variability and location

Now, checking for variability wrt location and scale for Computer Related Offenses (Sec 66 & Sec 66 B to E) in years 2016-21 using Levene's Test for variability and Kruskal Wallis Test for location

1.4.1 Levene Test

Null hypothesis, variances of all samples are equal against , variances of all samples are not equal

Test statistics,

$$W = \frac{(N-k)}{(k-1)} \cdot \frac{\sum_{i=1}^{k} N_i (Z_{i\cdot} - Z_{\cdot\cdot})^2}{\sum_{i=1}^{k} \sum_{j=1}^{N_i} (Z_{ij} - Z_{i\cdot})^2},$$

where, k: number of different groups to which the sampled cases belong. Ni: Number of elements in different groups. N: total number of cases in all groups

$$Z_{ij} = \left\{ \begin{array}{l} \left|Y_{ij} - \bar{Y}_i\right|, \bar{Y}_i \, is \, the \, mean \, of \, the \, i^{th} \, group \\ \left|Y_{ij} - \tilde{Y}_i\right|, \tilde{Y}_i \, is \, the \, median \, of \, the \, i^{th} \, group \end{array} \right\}$$

where, Y_{ij}: the value of jth case and ith group

Test is to reject null hypothesis for larger values of W

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value Pr(>F)
## group   5  0.6081 0.6938
##       162
```

Since p-value > 0.05 , for levene's test hence the test is not significant i.e. we failed to reject null hypothesis. Therefore there is no significant variability in the data through 2016-21

Now testing for equality in location using Kruskal Wallis Test

1.4.2 Kruskal Wallis Test

Null hypothesis, location of all populations is same against, location of any two populations differs

Test Statistics,

$$H = (N-1)\frac{\sum_{i=1}^{g} n_i(\bar{r}_{i\cdot} - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2}, \text{ where}$$

- N is the total number of observations across all groups
- g is the number of groups
- n_i is the number of observations in group i
- r_{ij} is the rank (among all observations) of observation j from group i
- $\bar{r}_{i\cdot} = \frac{\sum_{j=1}^{n_i} r_{ij}}{n_i}$ is the average rank of all observations in group i
- $\bar{r} = \frac{1}{2}(N+1)$ is the average of all the r_{ij}.

Test is to reject null hypothesis for large values OF H

```
##
##  Kruskal-Wallis rank sum test
##
## data:  cases by as.factor(year)
## Kruskal-Wallis chi-squared = 1.0211, df = 5, p-value = 0.9608
```

Since p-value > 0.05 , for Kruskal Wallis test, the result is not significant i.e. we failed to reject null hypothesis. Therefore there is no significant change in the locations throughout the years 2016-21

Hence, from both the tests,
Since there is no significant difference in location and variability in different years.We conclude that Computer Related Offenses (Sec 66 & Sec 66 B to E) are statistically same in terms of location and variability for different states.

1.5 State-wise comparison for Cases of Computer Related Offenses (Sec 66 & Sec 66 B to E)

1.5.1 Count of cases in each state

The figure 3 shows the highest number of cases being registered in Karnataka, followed by Uttar Pradesh and Assam. But one must note that the population in these states may vary. So to get a better comparison, rates can be calculated, which gives the cyber crime cases registered in the state per 1 lakh population. Mid year projected population for each state for the year 2016 - 2021 is available. Hence the Crime Rate can be calculated.

1.5.2 Crime rate in each state

The figure 4 shows the highest rate of cases being registered in Karnataka wrt 1 Lakh population in respective year, followed by Assam and UttarPradesh. But

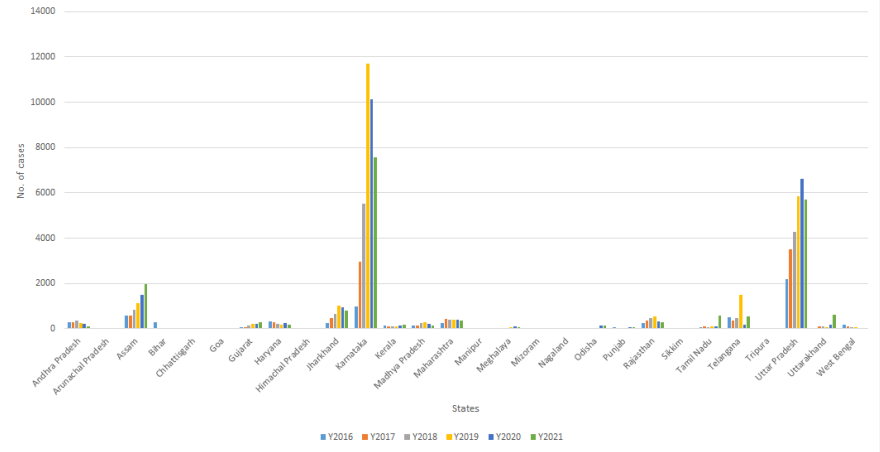

Figure 1.3: counts state-wise yearly CRO cases

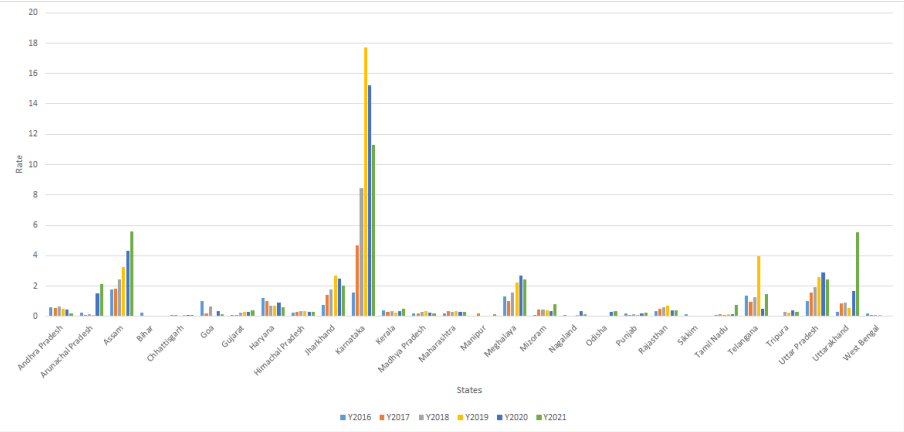

Figure 1.4: rates state-wise yearly CRO cases

one must note that the Uttrakhand had observed a peak in year 2021

The crime rate as well as counts of cyber crimes in Karnataka is surprisingly high. The pattern of crime rate wrt 1 Lakh population in respective year in Karnataka can be seen in figure 5.

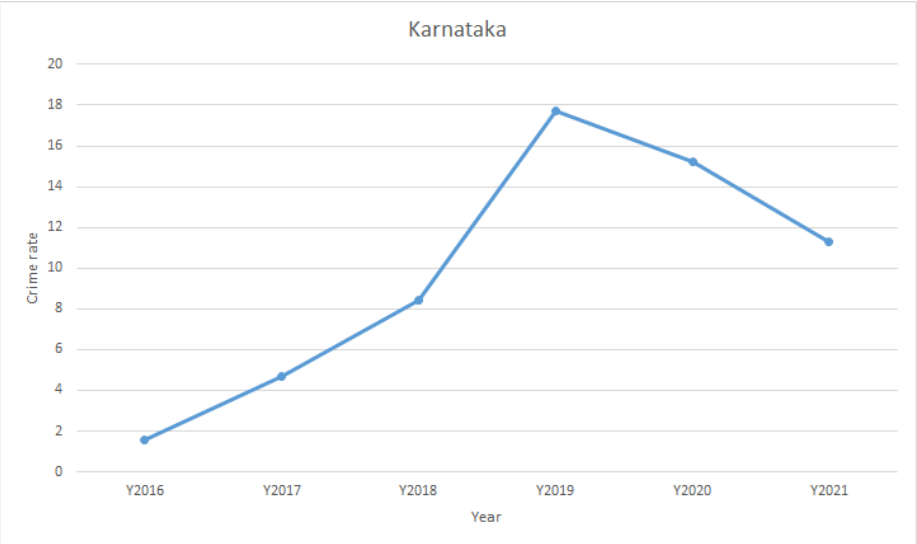

Figure 1.5: karnataka

1.5.3 Year wise increment of Computer Related Offenses (Sec 66 & Sec 66 B to E)

```
##      year       CRO  per_inc
## [1,] 2016 0.5351648       NA
## [2,] 2017 0.7844598  46.5828
## [3,] 2018 1.0685517  36.2150
## [4,] 2019 1.7652248  65.1979
## [5,] 2020 1.6198963  -8.2329
## [6,] 2021 1.4566373 -10.0784
```

The above data table shows the Computer Related Offenses got a peak in year 2019 when they increased by **65 %** as compared to previous year 2018. However a slight % decrease can be seen after that .