Awareness Of Cyber Crime Prevention Analyzed By The Nearest Neighbour Analysis In India

K.Ravichandran
Department of Visual Communication, Dr.MGR Educational and Research Institute, Chennai, Tamil Nadu, India

Abstract

The Development of Computer and mobile phone applications have increased cybercrime alarmingly in India. It takes several hours to detect and remove fraudulent information from the sharing of unauthorized systems. Otherwise, in cases of hacking information theft, they are intimidated by the internet. So we can protect our self with awareness of cyber security to avoid problems related to such cyber-rape incidents. For this a majority of the respondents that are 60 percent felt that only sometimes the media is giving cybercrime-related news, 25 percent said they did not find enough news related to cyber security while only 20 percent were of the opinion that cybercrime-related news appears in media very often. Of the remaining 20 percent, 5 percent were of the opinion that print media is creating more awareness and 15 percent said it is the electronic media that is creating people more aware of issues related to cybercrime. Quantitative analysis within the style of a survey was instrumental in collecting the information which is analyzed by the Nearest Neighbor Analysis. Along with this, measures being taken to control cybercrime also have also been discussed.

Keywords: *Cyber crime, Law, Internet Crime, Nearest Neighbor Analysis and Media Awareness.*

1. Introduction

India is the one of the populous countries and increase in crime rate is a cause of worry for the Indian economy and to the Indian government at large. Particularly computer based cyber crimes are also increasing and is a greater threat to the nation. Computers play a very significant role in modern life. A Number of person who has access to information on the computers, and most of the commercial and government information and unrestricted access to information present a real threat [1]. Protect the information structure from unauthorized dissemination. Moreover, unauthorized information needs great care and attention, and in general, events such as hacking, piracy, digital harassment, and digital blackmail must ensure the Digital Security Group [2]. Significantly, Website developers, Internet and networking software professionals are hardly able to control this menace. Such violations may debilitate a country's security and money related wellbeing issues encompassing these sorts of crimes have turned out to be prominent, especially those encompassing hacking, copyright encroachment, tyke erotic entertainment, and kid preparing [3].

2. Objective and Hypothesis of Study

The growing danger from crimes committed via the Internet, or against data on computers, is starting to claim attention of the world at large. This study investigates whether or not individuals would use the net to report crime. The main objective of the study is to find out media awareness among totally different respondents on the menace of cyber crime.

3. Methodology

Quantitative analysis inside the variety of a survey instrument has been used to accumulate the expertise and descriptive information to investigate the understanding. Quantitative analysis methodology was once used to support the genuine proven fact that the results of the survey have got to be part of the society who has entry to web. Because of the character of the evaluation, queries have been derived from the literature. These queries supplied a foundation for the analysis in an effort to get a transparent opinion about cyber crime among respondents and also to find out the sort of cyber crime so occurring within the latest days and what wishes to be completed to discontinue such crimes.

4. Result and Findings

The primary target respondents were working professionals who were aware of various computer related crimes and security issues within his/her organization. Typically, they included senior managers, IT administrators as well as IT consultants. Simple random sampling was the primary sampling method used when selecting the hundred samples for this survey.

A. Nearest Neighbor Analysis

The nearest Neighbor analysis gives us the exact information we need to accurately analyze the distance between each point and its nearest point. Subsequently, the results from a CSR (absolute spatial randomness) form are most accurately compared to the expected values for a random sample point. The CSR is generated by two assumptions: 1) all places are likely to be a case (event) and 2) all events are independent of each other.

B. Input

To enter the input data file, it must contain X, Y coordinates, and N rows of W values. It should be noted that all W values are equal to 1.

C. Analysis

The null hypothesis of CSR is all accurately tested using the Z statistic (standard normal variety). Positive score indicates dispersion or imbalance. This “CSR “Z statistic (standard normal type) is calculated very accurately using the formula below.

D. Formula:

a) The mean nearest neighbor distance

$$\bar{d} = \frac{\sum_{i=1}^{N} d_i}{N} \quad [1]$$

Where N is the number of points di is the nearest neighbor distance for point i.

b) The expected value of the nearest neighbor distance in a random pattern

$$E(d_i) = 0.5\sqrt{\frac{A}{N}} + \left(0.0514 + \frac{0.041}{\sqrt{N}}\right)\frac{B}{N} \quad [2]$$

Where A is the area and B is the length of the perimeter of the study area.

c) The variance

$$Var(\bar{d}) = 0.070\frac{A}{N^2} + 0.037B\sqrt{\frac{A}{N^5}} \quad [3]$$

d) The Z statistic

$$Z = \frac{\bar{d} - E(d_i)}{\sqrt{Var(\bar{d})}} \quad [4]$$

It should be noted that Equations [2] and [3] are based on Donnelly's (1978) [4] correction factor for calculating the boundary effect.

E. Output:

The following output files have listed: a) the input data file, b) the total number of points, c) the minimum and maximum of the X, and Y coordinates, d) the size of the study area, e) the observed mean nearest neighbor distance, g) the variance, and h) the Z statistic (standard normal variate).

F. Limitations:

Equations [2] and [3] cannot be used because we are the irregularly shaped study areas that need to be understood. In this project, the study area is a regular rectangle or square. The area (A) is computed (Xmax X Xmin) * (Ymax X Ymin), so we must realize that these are the boundaries of the study area here.

In method recognition, the K-Nearest Neighbor Algorithm (KNN) is a specially handled method for classifying objects based on the closest training examples in the feature space. KNN is a type of event-based learning, or lazy learning. We must understand that this function is only approximated locally and that all calculations are deferred until classification. Moreover, the KNN algorithm is not only the simplest of all machine learning algorithms, but also means that an object is classified by the majority of its neighbors, with its k assigned to the most common class among its Nearest Neighbors (k is a positive integer, usually small). If k = 1, it should be known that the object is assigned to its nearest neighbor[5].

In today's post, we will explore the application use of KNN for Internet users by creating prototypes of two new aspects of cybercrime awareness and how to prevent it. In this new venture, we want to raise awareness about cybercrime and how to prevent it through the media. Nowadays, we can see the level of awareness created by the media about the safe use of the Internet. And there are so many opportunities to learn how we can protect ourselves from cybercrime attacks. As we do this, we first start with the following data:

TABLE I
Table:1Case Processing Summary

		N	Percent
Sample	Training	76	76.0%
	Holdout	24	24.0%
Valid		100	100.0%
Excluded		0	
Total		100	

On Case Processing Summary (Table: 1), For Sample in Training is 76numbers and76.0%. In Holdout is 24 numbers and 24.0%.

G. Figures and Tables

Figure: 1 predictor Space

This map classifies our known patterns (in red) based on their predicted neighbors according to three predictor variables: cyber vandalism, victimization, and cyber terrorism. This is a 3D interactive chart that identifies 3 neighbors (K = 3) adjacent to our prototypes [6]. The next chart, called the Pierce chart, shows our prototypes with their nearest neighbors in the five input fields:

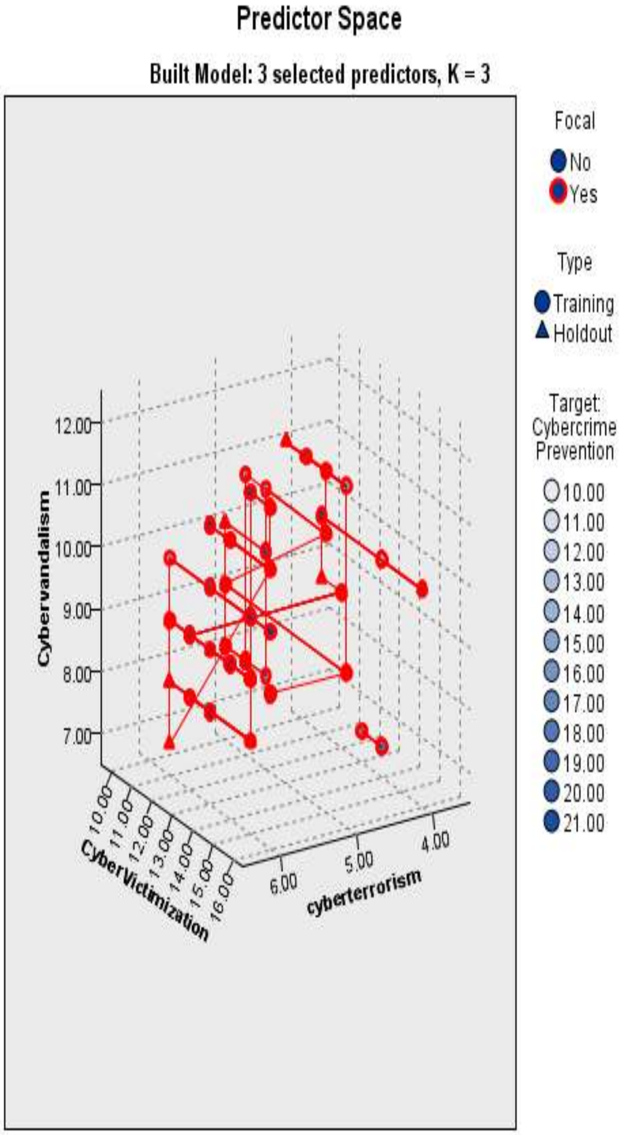
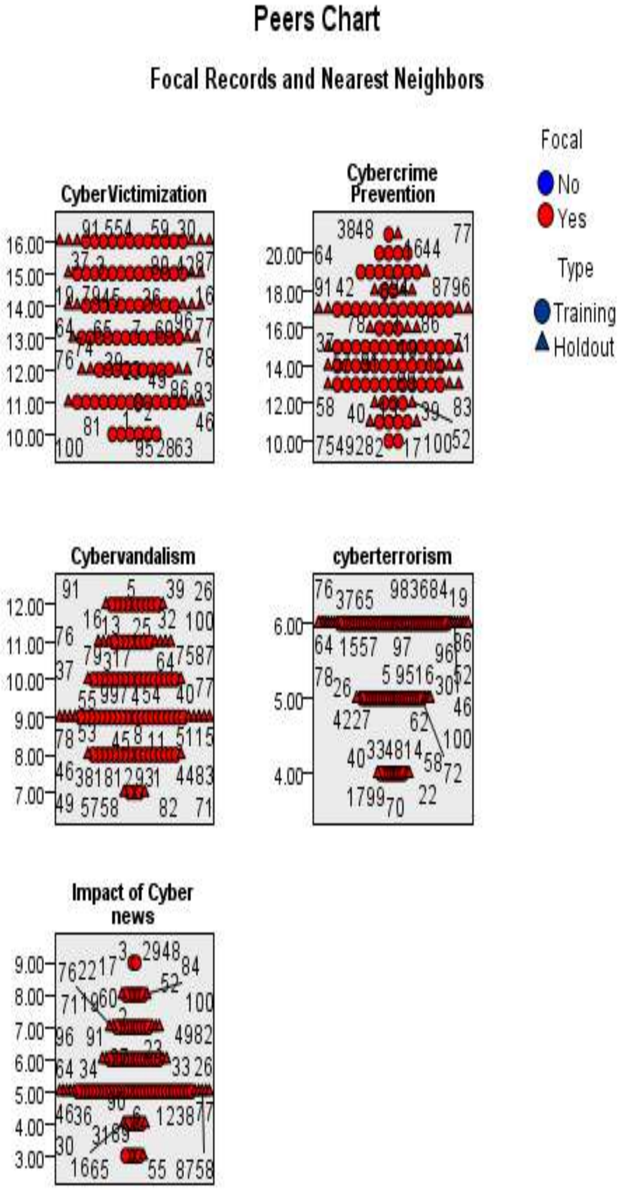
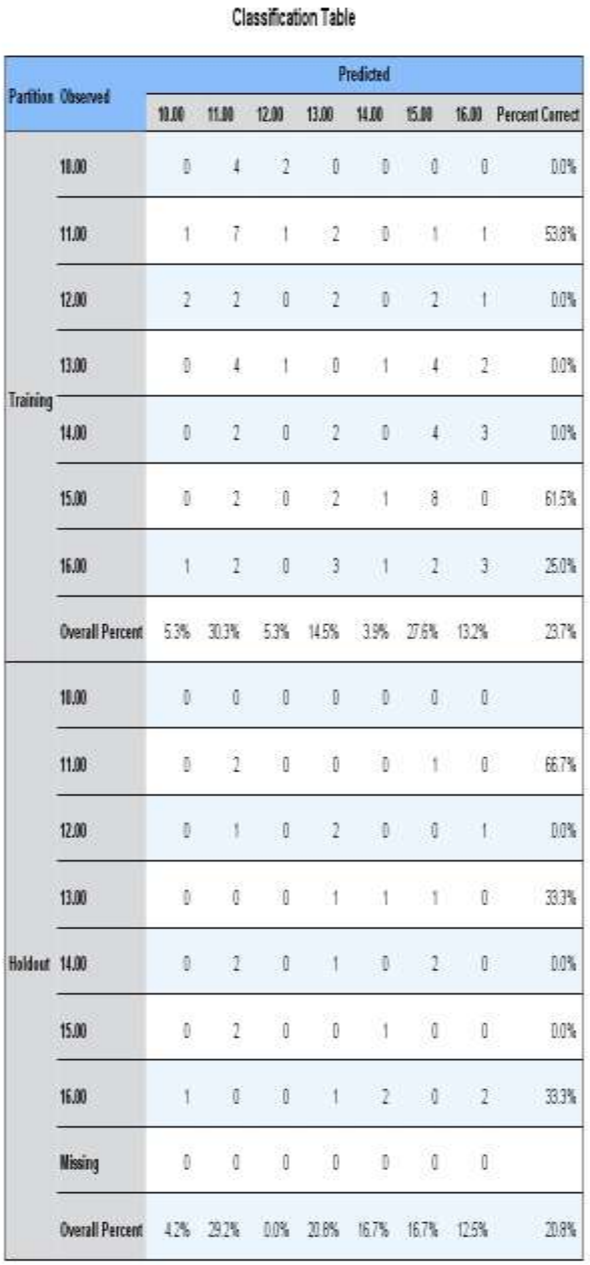


Figure: 2 Peers Chart



From this figure: 2 illustration, our prototypes take some form of cyber vandalism, cyber terrorism, victims, cybercrime prevention and cyber security awareness, whereas our prototypes are heavily cyber security compared to the current large scale Internet security service Is aware of the. Similar conclusions can be made in reviewing the other five entries. For example, it shows very large results even with very few samples [Annexure:1].

Next we review the k Nearest Neighbors and Distances chart:



Classification Table

Partition	Observed	Predicted							
		10.00	11.00	12.00	13.00	14.00	15.00	16.00	Percent Correct
Training	10.00	0	4	2	0	0	0	0	0.0%
	11.00	1	7	1	2	0	1	1	53.8%
	12.00	2	2	0	2	0	2	1	0.0%
	13.00	0	4	1	0	1	4	2	0.0%
	14.00	0	2	0	2	0	4	3	0.0%
	15.00	0	2	0	2	1	8	0	61.5%
	16.00	1	2	0	3	1	2	3	25.0%
	Overall Percent	5.3%	30.3%	5.3%	14.5%	3.9%	27.6%	13.2%	23.7%
Holdout	10.00	0	0	0	0	0	0	0	
	11.00	0	2	0	0	0	1	0	66.7%
	12.00	0	1	0	2	0	0	1	0.0%
	13.00	0	0	0	1	1	1	0	33.3%
	14.00	0	2	0	1	0	2	0	0.0%
	15.00	0	2	0	0	1	0	0	0.0%
	16.00	1	0	0	1	2	0	2	33.3%
	Missing	0	0	0	0	0	0	0	
	Overall Percent	4.2%	29.2%	0.0%	20.8%	16.7%	16.7%	12.5%	20.8%

Table: 2 Classification Table

In Classification Table:2 training, partision observed (11) online offer53%(overall 30%), partision observed (15) Web Security online offer53% (overall 27.6%), partision observed (16) Attachments25%(overall 13.3%).finally cyber crime prevention of web security awareness created by media.

5. Conclusion

Lack of awareness about the Internet as a tool to prevent crime was revealed. So there is no correlation between the level of media awareness of the respondents and the underestimation of the cyber crime threat to the community. It can be brought to the fact that it is a common misconception. One of the most important results is the potential exposure of the population in relation to the threats of cyber crime. Most of the respondents that is 80 per cent of them are of the opinion that social media is creating awareness among the public. How much are we safe, secure and reliable in this computer environment? Now, nothing has been confirmed. This is not just for our national security and for the Indian economy; this has prevented the development of different sciences and we are forced to consider the possibility of time. Therefore, it is important to recognize that dealing with cyber crimes is not an easy task without proper policy enforcement. However, these cyber crimes must be implemented by the government with proper rules and regulations so that people are safe in the cyber space and they will use the internet without any fear. To achieve this level, it is important that we protect the public from cyber crimes through public media awareness and raise awareness about how to deal with them safely. We can understand this survey more closely by analyzing Nearest Neighbor Analysis.

References

[1] Claycomb, W. R., & Nicoll, A. (2012, July). Insider threats to cloud computing: Directions for new research challenges. In 2012 IEEE 36th Annual Computer Software and Applications Conference (pp. 387-394). IEEE.

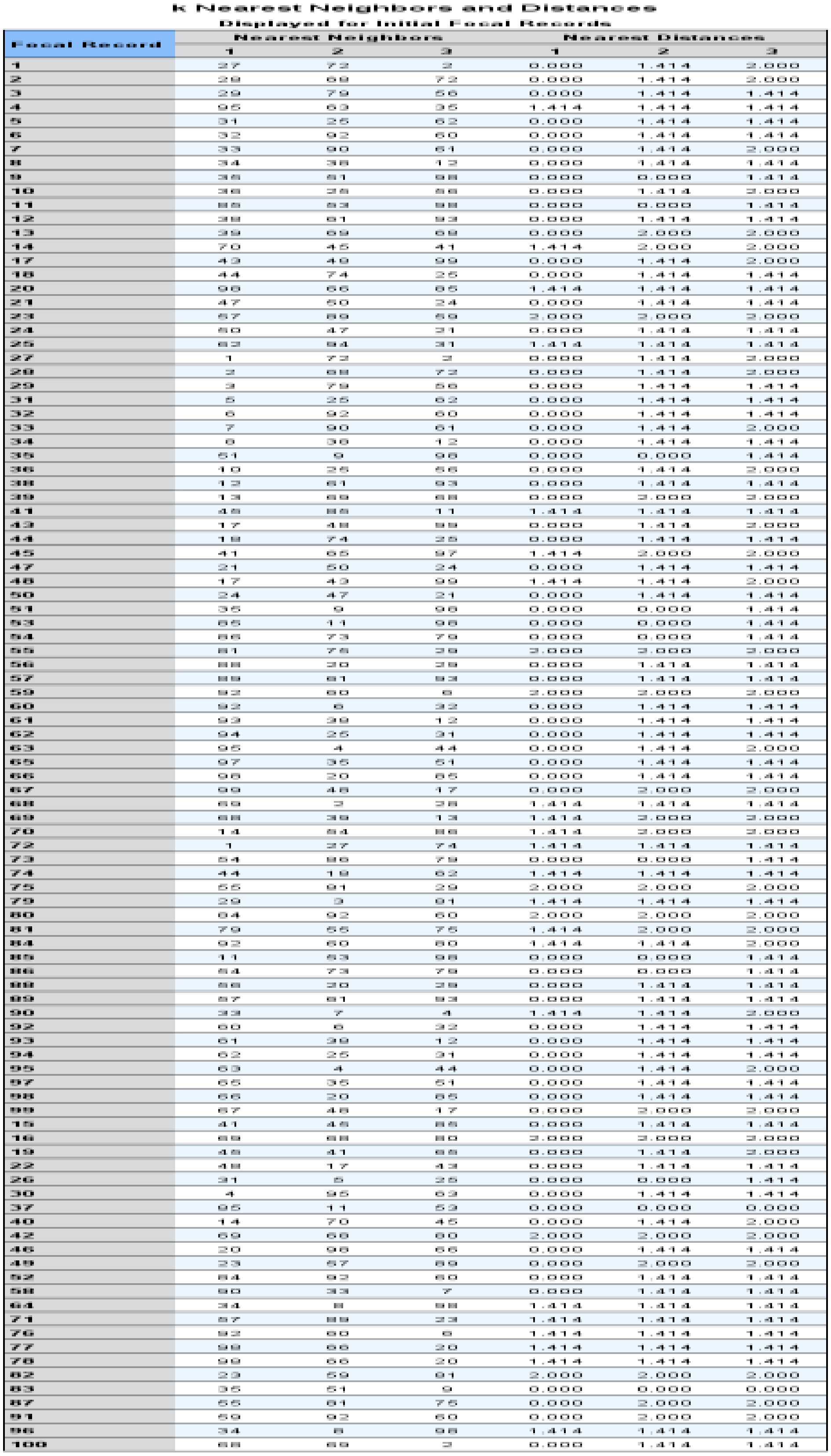
[2] Ravichandran, K., & Arulchelvan, S. (2017, February). Structural Equation Model Analyzed on Cyber Crime and Media Awareness in India. In 2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM) (pp. 141-146). IEEE.

[3] Hillyard, P., & Tombs, S. (2007). From 'crime'to social harm?. Crime, law and social change, 48(1-2), 9-25.

[4] Donnelly, H. (1978). On the wave equation asymptotics of a compact negatively curved surface. Inventiones mathematicae, 45(2), 115-137.

[5] Lakshmi, C. V., & Balakrishnan, J. R. (2012). Automatic accident detection via embedded GSM message interface with sensor technology. International Journal of Scientific and Research Publications, 2(4), 1.

[6] Konovalov, I. B. (2004). Local Optimization of Nearest Neighbors Based Models for Observed Time Series. Focus on Computer Science Research, 131.



k Nearest Neighbors and Distances
Displayed for Initial Focal Records

Focal Record	Nearest Neighbors			Nearest Distances		
	1	2	3	1	2	3
1	27	72	2	0.000	1.414	2.000
2	28	68	72	0.000	1.414	2.000
3	29	79	56	0.000	1.414	1.414
4	95	63	35	1.414	1.414	1.414
5	31	25	62	0.000	1.414	1.414
6	32	92	60	0.000	1.414	1.414
7	33	90	61	0.000	1.414	2.000
8	34	38	12	0.000	1.414	1.414
9	35	51	98	0.000	0.000	1.414
10	36	26	56	0.000	1.414	2.000
11	85	53	98	0.000	0.000	1.414
12	38	61	93	0.000	1.414	1.414
13	39	69	68	0.000	2.000	2.000
14	70	45	41	1.414	2.000	2.000
17	43	48	99	0.000	1.414	2.000
18	44	74	25	0.000	1.414	1.414
20	98	66	65	1.414	1.414	1.414
21	47	50	24	0.000	1.414	1.414
23	57	89	59	2.000	2.000	2.000
24	50	47	21	0.000	1.414	1.414
25	62	94	31	1.414	1.414	1.414
27	1	72	2	0.000	1.414	2.000
28	2	68	72	0.000	1.414	2.000
29	3	79	56	0.000	1.414	1.414
31	5	25	62	0.000	1.414	1.414
32	6	92	60	0.000	1.414	1.414
33	7	90	61	0.000	1.414	2.000
34	8	38	12	0.000	1.414	1.414
35	51	9	98	0.000	0.000	1.414
36	10	26	56	0.000	1.414	2.000
38	12	61	93	0.000	1.414	1.414
39	13	69	68	0.000	2.000	2.000
41	45	85	11	1.414	1.414	1.414
43	17	48	99	0.000	1.414	2.000
44	18	74	25	0.000	1.414	1.414
45	41	85	97	1.414	2.000	2.000
47	21	50	24	0.000	1.414	1.414
48	17	43	99	1.414	1.414	2.000
50	24	47	21	0.000	1.414	1.414
51	35	9	98	0.000	0.000	1.414
53	85	11	98	0.000	0.000	1.414
54	86	73	79	0.000	0.000	1.414
55	81	75	29	2.000	2.000	2.000
56	88	20	29	0.000	1.414	1.414
57	89	61	93	0.000	1.414	1.414
59	92	60	6	2.000	2.000	2.000
60	92	6	32	0.000	1.414	1.414
61	93	38	12	0.000	1.414	1.414
62	94	25	31	0.000	1.414	1.414
63	95	4	44	0.000	1.414	2.000
65	97	35	51	0.000	1.414	1.414
66	98	20	65	0.000	1.414	1.414
67	99	48	17	0.000	2.000	2.000
68	69	2	28	1.414	1.414	1.414
69	68	39	13	1.414	2.000	2.000
70	14	84	88	1.414	2.000	2.000
72	1	27	74	1.414	1.414	1.414
73	54	86	79	0.000	0.000	1.414
74	44	18	62	1.414	1.414	1.414
75	55	81	29	2.000	2.000	2.000
79	29	3	81	1.414	1.414	1.414
80	84	92	60	2.000	2.000	2.000
81	79	55	75	1.414	2.000	2.000
84	92	60	80	1.414	1.414	2.000
85	11	53	98	0.000	0.000	1.414
86	54	73	79	0.000	0.000	1.414
88	56	20	29	0.000	1.414	1.414
89	57	61	93	0.000	1.414	1.414
90	33	7	4	1.414	1.414	2.000
92	60	6	32	0.000	1.414	1.414
93	61	38	12	0.000	1.414	1.414
94	62	25	31	0.000	1.414	1.414
95	63	4	44	0.000	1.414	2.000
97	65	35	51	0.000	1.414	1.414
98	66	20	65	0.000	1.414	1.414
99	67	48	17	0.000	2.000	2.000
15	41	45	85	0.000	1.414	1.414
16	69	68	80	2.000	2.000	2.000
19	45	41	85	0.000	1.414	2.000
22	48	17	43	0.000	1.414	1.414
26	31	5	25	0.000	0.000	1.414
30	4	95	63	0.000	1.414	1.414
37	85	11	53	0.000	0.000	0.000
40	14	70	45	0.000	1.414	2.000
42	69	68	80	2.000	2.000	2.000
46	20	98	66	0.000	1.414	1.414
49	23	57	89	0.000	2.000	2.000
52	84	92	60	0.000	1.414	1.414
58	90	33	7	0.000	1.414	1.414
64	34	8	98	1.414	1.414	1.414
71	57	89	23	1.414	1.414	1.414
76	92	60	6	1.414	1.414	1.414
77	98	66	20	1.414	1.414	1.414
78	99	66	20	1.414	1.414	1.414
82	23	59	91	2.000	2.000	2.000
83	35	51	9	0.000	0.000	0.000
87	55	81	75	0.000	2.000	2.000
91	59	92	60	0.000	2.000	2.000
96	34	8	98	1.414	1.414	1.414
100	68	69	2	0.000	1.414	1.414