

Article

Analyzing the Problems of a District-Based Administration Using Monte Carlo Simulation: The Case of Sex Offender Notifications in Korea

Hyemin Kim, Suyun Lee and Chulmin Jun * 

Department of Geoinformatics, University of Seoul, 163 Seoulsiripdae-ro, Dongdaemun-gu, Seoul 02504, Republic of Korea; kimhm77@uos.ac.kr (H.K.); stbry27@uos.ac.kr (S.L.)

* Correspondence: cmjun@uos.ac.kr

Abstract: The problems of administrations based simply on administrative units that do not consider the operational purposes of the system have been consistently discussed. For example, in the Republic of Korea, sex offenders' information is distributed via physical mail only in a few regions, a practice that is too rigidly based on the boundaries of the administrative 'Dong' of the offender's residence. This implies that citizens in an adjacent building will not be notified if their Dong is different. Therefore, this study analyzed the problems of an administrative system that does not consider its realistic scope by using the case study of sex offender notifications. By expanding the distance from children and youth grids, we ascertained the extent of the problems with sex offender notifications. Additionally, to determine whether these problems have occurred by chance at a specific point in time or if there has been a fundamental limitation in the system, the Monte Carlo simulation was applied to compare the actual and random data of residences.

Keywords: administrative district; sex offender information notices; Monte Carlo simulation



Citation: Kim, H.; Lee, S.; Jun, C. Analyzing the Problems of a District-Based Administration Using Monte Carlo Simulation: The Case of Sex Offender Notifications in Korea. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 177. <https://doi.org/10.3390/ijgi13060177>

Academic Editors: Wolfgang Kainz and Jamal Jokar Arsanjani

Received: 13 March 2024

Revised: 16 May 2024

Accepted: 26 May 2024

Published: 29 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the Republic of Korea, for effective administration, cities are divided into administrative districts such as Gu and Dong—for example, Seoul has 25 Gus and 426 Dongs. However, administrations based on administrative units that do not consider the operational purposes of the system continues to be an issue [1–3]. For example, the nuclear safety (radioactive leak) crisis response practice manual specifies that in the case of a fire or radioactive leak, the responsible fire department will be determined based on the administrative district. Thus, during an actual incident, even though a 911 safety center was located 2 km away, a center 11 km away was designated as the fire department in charge. Consequently, the response time took more than five times as long [4–6]. It indicated that the designation of centers based on administrative districts was not suitable for safety fields requiring a rapid response, and therefore, the designation of fire departments was changed to one based on distance.

The need to consider the distance standards when setting the scope can also be identified in the issue of sex offender notifications. In the Republic of Korea, sex offenders' information is distributed via physical mail in only a few regions. Currently, the information of sex offenders is distributed to households and institutions that protect children and youth in the administrative Dong where the offender resides. Sex offenders are more likely to commit crimes in familiar places than unfamiliar places [7], and more than half of sex crimes against minors occur near the offender's residence [8]. However, the administrative Dong-based information notification system has a limitation in that the information is not delivered to residents who actually need the information. This is because even if the offenders reside in an adjacent building, their information cannot be received by residents unless the offender is residing within the same administrative Dong.

An example of this is shown in Figure 1. The thick line represents the boundary of the administrative district, the circle represents a residence with children and youth, and the triangle represents a residence of an offender. As shown in Figure 1b, children and youth can receive information about offender X2 who is relatively far away, but as seen in Figure 1a, the information on offender X1, who is located very close by, cannot be received by the children and youth because the offender's residence is not in the same administrative district. To provide practical information, it is necessary to revise the notification system to one that is based on distance from the residence of the sex offender. Therefore, an analysis of the problems caused by the current notification system can guide the basic research for effective policy improvements.

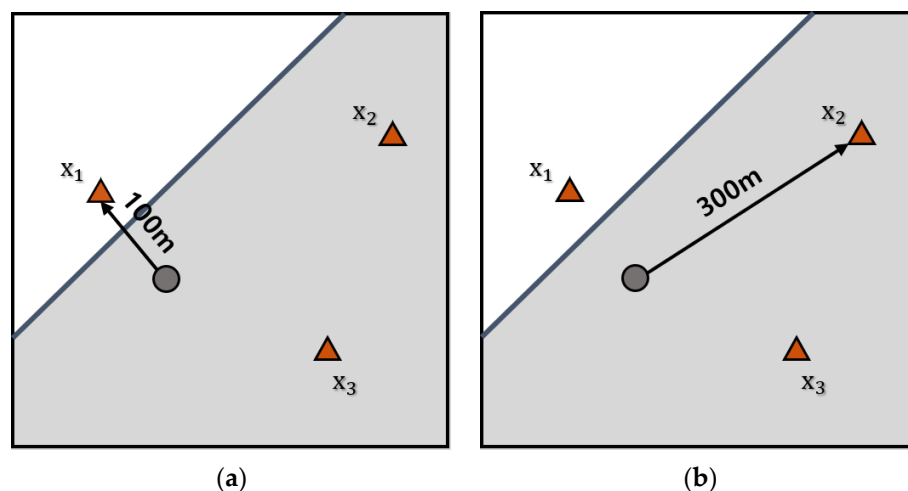


Figure 1. Notification status of offenders: (a) different districts, (b) same districts.

After analyzing the problems, we attempt to validate and compare the differences between the random results and actual outcomes through Monte Carlo simulations, further confirming the existence and inevitability of the problems with a notification system based on administrative boundaries. The Monte Carlo simulation has been widely used in various fields for problem solving. This method generates random numbers for situations where direct experimentation is difficult, repeating a given trial many times and using the probabilities obtained as a guide [9–12]. This simulation includes a comprehensive mechanism for obtaining statistical values for the data via repeated performance and, thereafter, performing statistical significance tests or calculating confidence intervals for the observed statistical values [13,14]. Previous research has also applied Monte Carlo simulations to perform statistical significance tests for complex problems [15,16]. Particularly, owing to the mathematical complexity of spatial point processes, most statistical tests for spatial point patterns have been performed using Monte Carlo procedures [17,18]. As the residence data of sex offenders used in this study were obtained at a specific point in time, the study results are dependent on the sex offenders' residence. Therefore, by comparing the results of a study analysis with the results of many random repetitions, we can be sure that the research results were not obtained by chance at a specific point in time.

The purpose of this study is to analyze the limitations of the administrative district unit system by using the sex-offender-notification issue as a case study. To this end, by increasing the distance centered on the children and youth grid, we derive the number of sex offenders living in the nearby area, the ratio, and the distance from which sex offenders' information is not delivered. Additionally, to verify that these problems do not occur only in certain residence types, the Monte Carlo simulation was applied to compare the degree of sex offender notification problems in the residence distribution of this study and random types.

2. Data and Study Area

The 2023 Korean census reveals that Seoul—the capital of Korea—has high population density and a population of approximately 9.65 million people. Seoul is also the region with the second highest number of sex offenders. Considering Seoul's high population density and number of sex offenders, inappropriate notification methods could have a direct impact on residents' safety and cause a lot of inconvenience. The data used for the analysis are the residence points of sex offenders in Seoul as of 1 October 2022, published on the Sex Offender Notification site (<https://www.sexoffender.go.kr/> (accessed on 25 May 2024)) and population data on a 100 m grid provided by the National Geographic Information Institute (<https://map.ngii.go.kr/ms/map/NlipMap.do> (accessed on 25 May 2024)). Currently, the information of sex offenders is distributed to household and institutions that protect children and youth in the administrative Dong where the offender resides. Therefore, among the population provided, data on pre-, elementary, middle, and high school students were combined and processed into a new set of child and youth population data. Grid center points were used to calculate the distance between the child and youth population and the offenders' residences. Additionally, while the sex offenders' residence data provide accurate addresses, the retrieved data have been anonymized and actual residences have not been visualized due to privacy issues.

3. Methodology

3.1. Number of Sex Offenders by Distance

To determine how many sex offenders live near child- and youth-care households, the number of sex offenders living within the area was calculated by expanding the distance in 100 m increments from 100 m to 1 km based on the center point of each grid. In grid A in Figure 2, if the circle denotes the center point of the grid and triangle denotes a sex offender's residence, then it can be seen that three offenders live within 200 m of grid A. In this way, the number of nearby offenders was calculated for all the grids in Seoul. As the significance of the results varies depending on the number of children and youth within a grid, the average number of sex offenders by distance is calculated by applying the number of children and youth residing in each grid as a weight. The specific formula used in this study is notated as follows:

$$An(d) = \frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n p_i} \quad (1)$$

where $An(d)$ represents the average number of sex offenders within distance d , n represents the total number of cells, p_i represents the number of children and youth in cell i , and x_i represents the number of sex offenders within distance d from cell i .

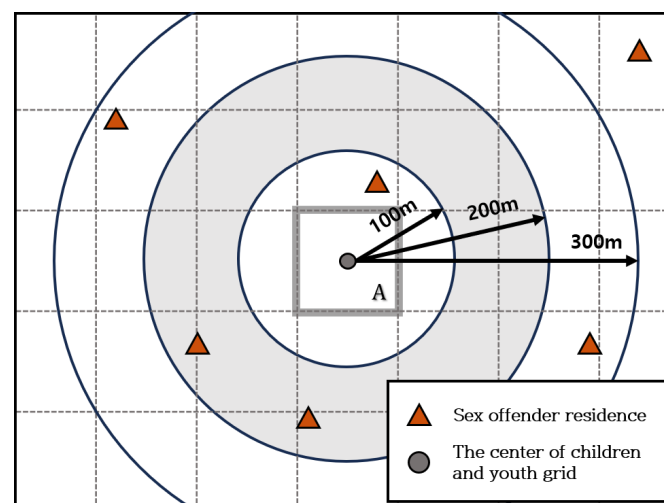


Figure 2. Average number of sex offenders living nearby.

3.2. *k*-Nearest Neighbor

To ascertain the distance between child- and youth-care households and sex offenders living nearby, the distance was calculated from the closest sex offender to the 10th closest sex offender based on the center point of each grid (Figure 3). This was calculated for all the grids in Seoul to derive the average distance to the *k*th closest sex offender. As with the previous analysis, the number of children and youth in the grid was weighted by for each grid. The specific formula is as follows:

$$Ad(k) = \frac{\sum_{i=1}^n p_i y_i}{\sum_{i=1}^n p_i} \quad (2)$$

where $Ad(k)$ denotes the average distance to the *k*th nearest sex offenders, *n* denotes the total number of cells, p_i denotes the number of children and youth in cell *i*, and y_i denotes the distance to the *k*th nearest sex offenders from cell *i*.

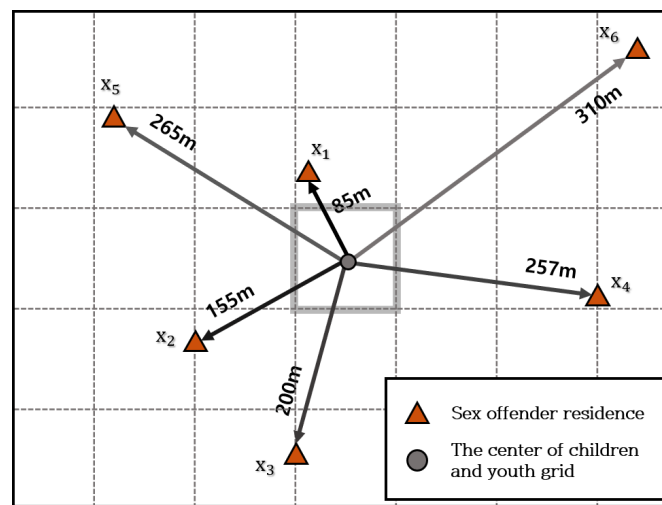


Figure 3. Average distance from sex offenders living nearby.

3.3. Vulnerability Statistics 1: Notification Rate

To ascertain the rate of notifications for information about sex offenders living nearby, the distance is expanded in 100 m increments from 100 m to 1 km based on the center point of each grid, and the proportion of sex offenders residing within the area whose information was distributed is calculated. If the circle denotes the center of children and youth grid, the triangle denotes the residence of a sex offenders, and the thick straight line denotes the Dong (district) boundary, then grid A in Figure 4 shows a total of three sex offenders within 500 m—one of these is a known offender and two are not known because they do not reside in the same administrative district. Therefore, the notification rate of the grid can be calculated as 33.3%. This was calculated for all the grids in Seoul, weighted by the number of children and youth in the grid. The equation for calculating the average sex offender notification rate by distance is given as:

$$Ap(d) = \frac{\sum_{i=1}^n p_i z_i}{\sum_{i=1}^n p_i}, \quad z_i = \frac{a_i}{x_i} \quad (3)$$

where $Ap(d)$ is the average percentage of notified sex offenders within distance *d*, *n* is the total number of cells, p_i is the number of children and youth in cell *i*, a_i is the number of sex offenders residing in the same Dong as cell *i* within distance *d*, and x_i is the number of sex offenders within distance *d* from cell *i*.

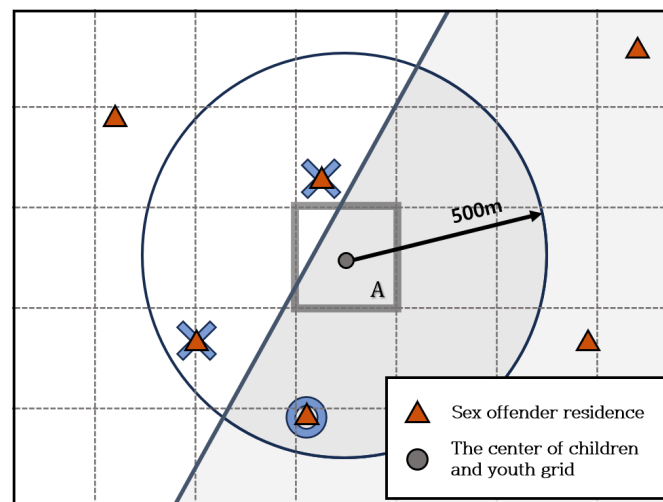


Figure 4. Average percentage of sex offenders living nearby whose information has been distributed.

3.4. Vulnerability Statistics 2: Distance

To ascertain the distance between child- and youth-care households and sex offenders for whom information is not distributed, the distance from each grid center to the nearest offender residing in a different administrative district was calculated. Figure 5 presents an example, and like the previous indicators, the number of children and youth in the grid was applied as a weight. This was calculated for all grids to verify the distribution, and the average value was derived. Additionally, because such notification problems occur more often as sex offenders live closer to administrative district boundaries, the average distance between a sex offender's residence and administrative district boundaries was calculated, which is shown in Figure 6.

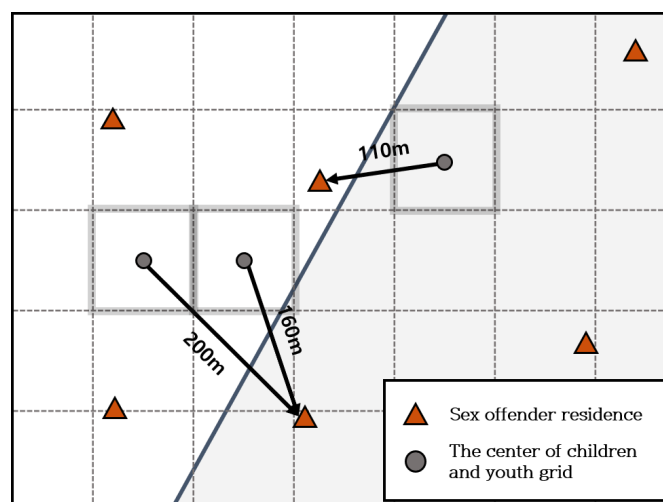


Figure 5. Distance from sex offenders whose information has not been distributed.

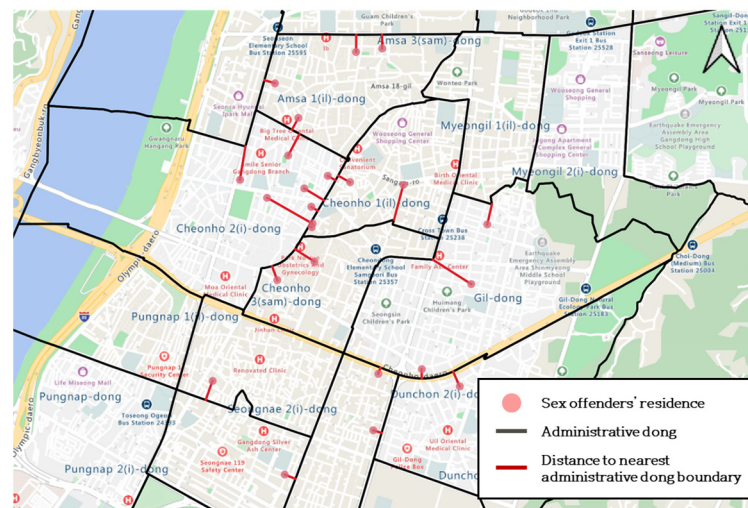


Figure 6. Average distance between sex offenders and the nearest administrative Dong boundary.

3.5. Monte Carlo Simulation

This study was analyzed based on the actual distribution of citizen residence and sex offender residence in Seoul, Republic of Korea. An analysis based on the distribution of residences at such a specific point in time may raise questions about whether the problems arose due to the incidental distribution of residences at that time. Thus, to confirm that the problem of sex offender notifications does not occur only in the specific residential distribution of offenders, the model was built to randomly generate residences and calculate the rate of notifications and the distance from sex offenders whose information is not distributed in the generated residences. The number of sex offenders generated in the simulation was fixed as equal to the actual number of sex offenders used in this study. The experiment of randomly generating sex offender residences in habitable areas excluding mountains, rivers, and industrial complexes in Seoul and calculating statistical values was repeated 999 times [19]. The Monte Carlo simulations applied to each were programmed using R 4.2.2 x64, and the aggregated values were ranked from the lowest to highest. To compare the results of the randomly generated residence-based analysis with the actual residence-based analysis, a 95% confidence interval was derived by excluding the bottom-50 observations among the sorted statistics [15,16]. If the results of the actual residence-based analysis were included in the confidence interval, it was determined that the level of notification problems occurring in the two residence types was similar. Additionally, it was considered that problems similar to the analysis results would occur in any type of residence. In other words, the Monte Carlo simulation and testing process can be summarized as follows:

- Step 1. Generate sex offender residence points randomly in the residential areas within Seoul.
- Step 2. For the generated random residence points, calculate the notification rate (or distance to a sex offender whose information has not been distributed).
- Step 3. Repeat steps 1 and 2 a total of 999 times.
- Step 4. Sort the calculated statistics from all random samples from the lowest to highest.
- Step 5. Derive a 95% confidence interval for each model by removing the bottom-50 observations.
- Step 6. If the actual sex-offender-residence-based analysis results fall within the 95% confidence interval of the randomly generated residence-based results, the level of notification problems is considered similar for the two residence types.

4. Results

4.1. Number of Sex Offenders by Distance

To ascertain how many sex offenders resided near child- and youth-care households, the average number of sex offenders residing within the area in 100 m increments from 100 m to 1 km was calculated. The analysis revealed that the average number of sex offenders residing within 1 km—alternatively, within a 15 min walk—from child- and youth-care households is three (Figure 7).

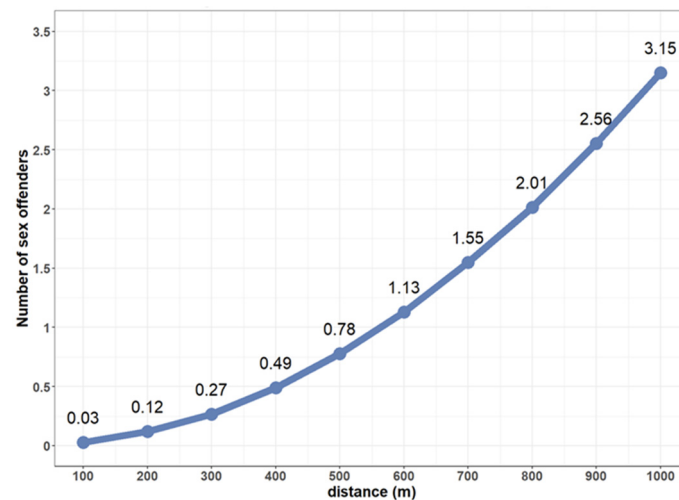


Figure 7. Result of the average number of sex offenders within distance d .

4.2. k -Nearest Neighbor

Figure 8 shows the average distance from the nearest sex offender to the 10th closest sex offender based on the center point of each grid to ascertain the distance between child- and youth-care households and sex offenders residing nearby. The analysis shows that the average distance from children and youth to the nearest sex offender is approximately 500 m, which implies that sex offenders reside within 500 m of their location, which is generally considered walking distance in Korea.

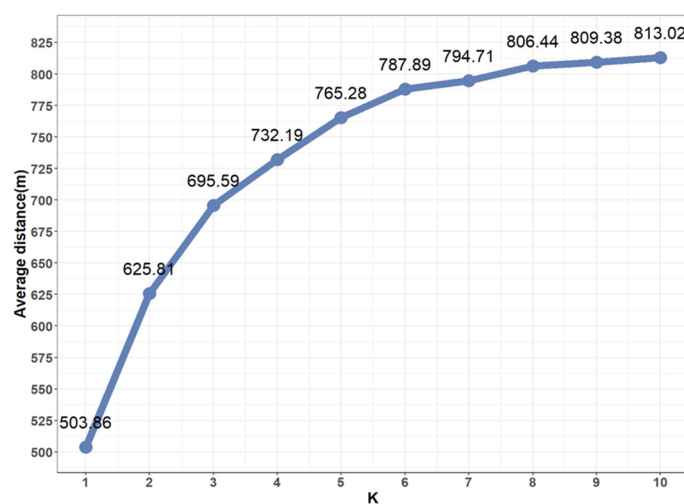


Figure 8. Result of the average distance to the k th nearest sex offenders.

4.3. Vulnerability Statistics 1: Notification Rate

To ascertain the rate of notifications for information about sex offenders residing nearby, we calculate the average rate of sex offenders residing in the area whose information

is distributed. As shown in Figure 9, by expanding the distance in 100 m increments from 100 m to 1 km based on the center of each children and youth grid, the average rate of notifications for information about sex offenders based on a 500 m distance is approximately 0.56. This implies that, for example, if there are four sex offenders residing within 500 m of children and youth, approximately half of them (two people) will not have their information distributed.

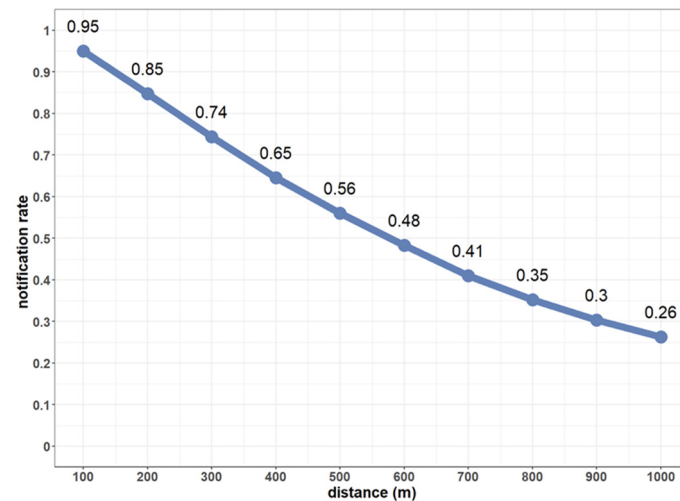


Figure 9. Result of average percentage of sex offenders whose information has been distributed within d distance.

4.4. Vulnerability Statistics 2: Distance

To ascertain how close child- and youth-care households are to sex offenders whose information has not been distributed, the distance to the nearest offender residing in another administrative district was calculated based on the center of each grid and displayed as a distribution. The analysis revealed approximately 255,000 children and youth households within 500 m of sex offenders whose information has not been distributed, as shown in Figure 10. This implies that approximately 28% of child- and youth-care households are not notified about sex offenders residing within 500 m.

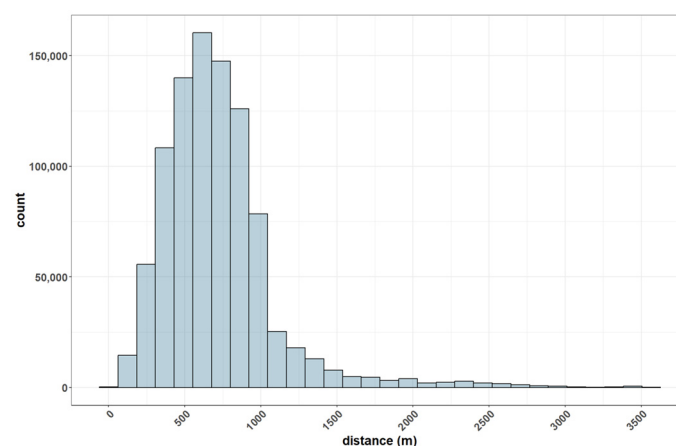


Figure 10. Distribution of distance to sex offenders whose information has not been distributed.

As a result of verifying how close the offender (whose information was not distributed) was to each child- and youth-care household, the number of such offenders who are the closest to the child- and youth-care households was the largest at 547,044; the next largest was the second closest (Figure 11).

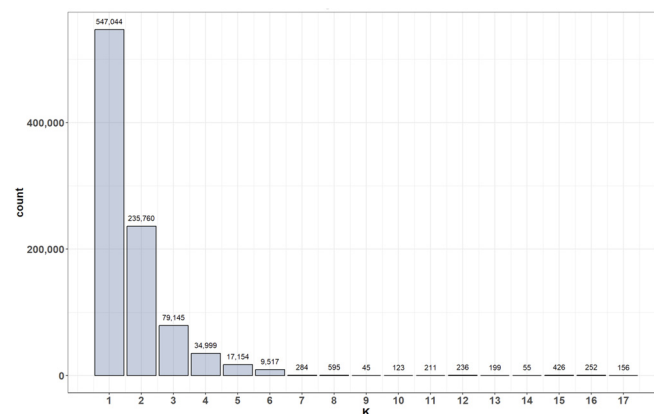


Figure 11. Close ranking of sex offenders whose information has not been distributed.

These notification problems are more likely to occur when sex offenders reside closer to administrative district boundaries. Therefore, the average distance between the sex offender's residence and administrative district boundaries was calculated. The analysis revealed that this was approximately 143 m—much closer than the 500 m that is commonly referred to as walking distance—and an average of 2 min walk. This implies that a sex offender could be residing at a walking distance from multiple districts. Therefore, the current notification system requires amendment as it does not deliver a sex offender's information who resides in a different administrative district.

4.5. Monte Carlo Simulation

The Monte Carlo simulation was performed to confirm that the notification problem does not occur only in the specific residence type of sex offenders. After 999 calculations of the notification rate for randomly generated residence points performed in the same way as the section "Vulnerability Statistics 1", the rate of notifications for information about sex offenders within 500 m of children and youth was a minimum of 0.58 and a maximum of 0.66. Although the notification rate in this study was 0.56, the actual residence-based analysis results did not fall within the 95% confidence interval (blue dash line in Figure 12) of the randomly generated residence-based results. Figure 12 shows that the random residence type of sex offenders had a higher notification rate than the residency type used in this study. However, even in the random residency type, if there were three offenders within 500 m, at least one of them would not have their information distributed.

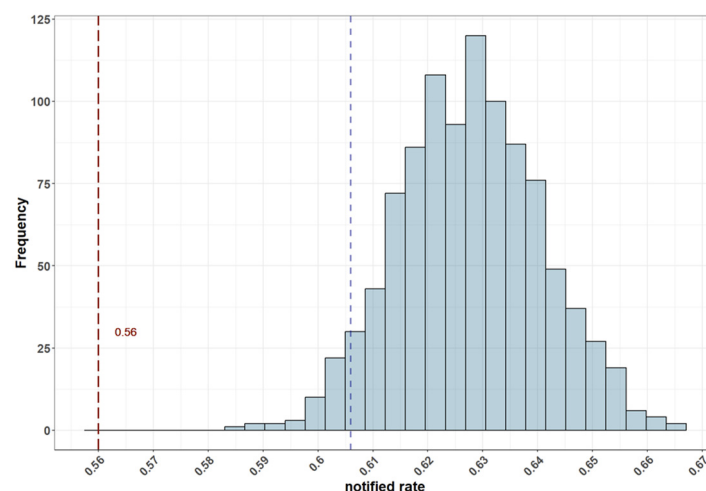


Figure 12. Results of Monte Carlo Simulation 1 (notification rate).

Similarly, the residences of sex offenders were randomly generated, and the distance from sex offenders whose information was not distributed was calculated 999 times, like the section “Vulnerability Statistics 2”. The simulation results show that the average distance from such a sex offender is a minimum of 678 m and a maximum of 756 m for random sex offender residences (Figure 13). The average distance from such sex offenders in this study was 716.86 m. The actual residence-based analysis results fall within the 95% confidence interval (blue dash line in Figure 13) of the randomly generated residence-based analysis results. It was confirmed that even in random residence types, child- and youth-care households are similarly close to such sex offenders. This implies that the problem of distributing sex offender information is not a coincidence of residence at a particular point in time. Rather, it is a fundamental limitation of the administrative-district-based system. Therefore, these problems cannot be avoided at any residence.

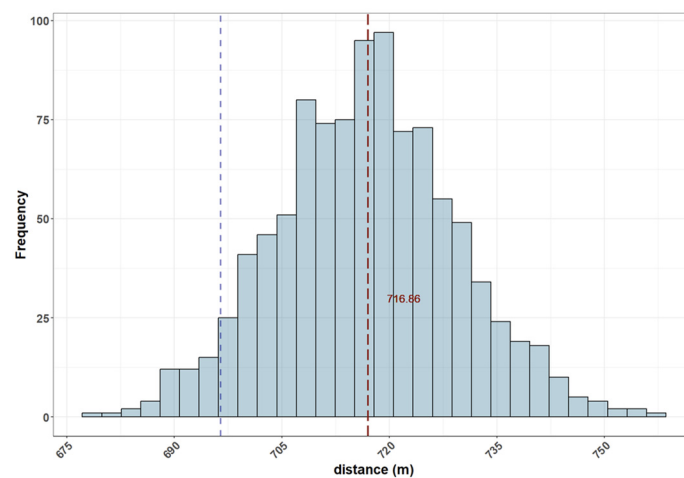


Figure 13. Results of Monte Carlo Simulation 2 (distance).

5. Discussion

As confirmed by the analysis results, the persistent occurrence of the same issues in arbitrary residential types implies that the administrative-Dong-based notification system cannot properly provide information to those who need that information. This notification system is more appropriate for delivering notifications based on distance rather than administrative districts determined for convenience. Therefore, we calculated the distance that can provide information at the same cost as the current notification system. The analysis was conducted on the assumption that if notifications were to be made to the same number of people as those notified under the existing system, the cost of notifications would be the same. As a result, it was found that approximately 600 m would yield the same cost as the existing notification system. To determine the appropriate notification standards, various factors must be considered, including legal conditions, average walking distance, and the budget available for administrative processing. In future research, it is expected that by taking multiple factors into account, a more appropriate distance standard can be established to provide notifications to those who need information.

6. Conclusions

This study analyzed the problem of administrations based on fixed areas, such as administrative districts, without considering the spatial scope of the system’s operational purpose. The case study of sex offender notifications was used as an example. The analysis confirmed that sex offender information was not properly delivered to households that actually needed the information. To obtain the intended effects of the system’s operational purpose, a spatial perspective must be considered while setting the scope. The lack of notifications regarding sex offender information could adversely impact residents’ safety; therefore, individuals who need this information must be properly notified.

The analysis considered only the limited variables of child- and youth-care households, among those in need of information via notifications. Therefore, if data on child- and youth-care institutions and the population using the facilities are to be used for analysis, it is believed that the extent of problems occurring could be confirmed in greater detail. However, rather than quantifying the specific problem, it is necessary to first analyze how improvements can be made to the notification system—that takes into account one’s distance from an offender’s residence—and review the spatial scope of the existing administrative set up. The Republic of Korea is currently considering a “Korean’s Jessica Law”. Jessica’s Law in the USA—which prohibits sex offenders from residing within a certain distance of child- and youth-care households—may not apply to high-density areas like Seoul. It is expected that this study will be used as basic data to improve sex-offender-information notification systems and the simple administrative-district-based system.

Author Contributions: Hyemin Kim and Chulmin Jun conceived the study; Hyemin Kim, Suyun Lee and Chulmin Jun designed the research; Hyemin Kim and Suyun Lee acquired the data; Hyemin Kim and Suyun Lee conducted the investigation; Hyemin Kim performed the analyses; Hyemin Kim performed the visualization and validation and wrote the manuscript; Hyemin Kim, Chulmin Jun and Suyun Lee read and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the 2023 Research Fund of the University of Seoul.

Data Availability Statement: All relevant data are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Kim, K.W. “Nothing Has Changed Since the ASF Outbreak” ... Administrative Convenience Standards Still Remain. *Korea Farmers and Fishermen Newspaper*. 2021. Available online: <https://www.agrinet.co.kr/news/articleView.html?idxno=303781> (accessed on 25 May 2024).
- Lee, I.H. There Are No Cases of Quarantine Policies at the Administrative District Unit. *Livestock Newspaper*. 2019. Available online: <http://chuksannews.co.kr/mobile/article.html?no=215763> (accessed on 25 May 2024).
- Yoon, S.J. Association of Swine Veterinarians ‘ASF Stamping out and Movement Restriction Period Must Be Reduced’. *DailyVET*. 2022. Available online: <https://www.dailyvet.co.kr/news/policy/164980> (accessed on 25 May 2024).
- Park, J.H. [National Fire Service Inspector General] Jaeho Park “The Fire Department Should Be Designated as a Distance, Not an Administrative District”. *Fire Prevention News*. 2020. Available online: <https://www.fpn119.co.kr/144789> (accessed on 25 May 2024).
- Song, B.H. Nuclear Disaster Manual, in Case of an Accident, a ‘Further’ Fire Department Responds. *Energy Daily*. 2020. Available online: <http://www.energydaily.co.kr/news/articleView.html?idxno=112535> (accessed on 25 May 2024).
- Jo, D.E. ‘Nuclear Disaster Manual’ That Require the Fire Departments Further away to Respond Should Be Reorganized. *Today Energy*. 2020. Available online: <http://www.todayenergy.kr/news/articleView.html?idxno=229657> (accessed on 25 May 2024).
- Kim, J.Y.; Jung, S. Geographical Profiling of Serial Rape in Korea. *Korean J. Public Saf. Crim. Justice* **2011**, *20*, 37–58.
- Gu, M.J. [Exclusive] 44% of Sex Offenders in Seoul Live in Child Protection Areas. *Weekly Newsmagazine*. 2023. Available online: http://www.sisajournal.com/news/articleView.html?idxno=253854#google_vignette (accessed on 25 May 2024).
- Ahn, Y.H. A Study on the Probabilistic Decision Model of the Overseas Railway Investment Project Applying Monte Carlo Simulation. Ph.D. Thesis, Woosong University, Daejeon, Republic of Korea, 2019; pp. 28–32.
- Heo, C.K.; Jung, C.K. The strategic decision-making of volleyball game using Monte-Carlo Simulation. *Korean Soc. Sports Sci.* **2023**, *32*, 739–749.
- Kim, S.G. A Study on Teaching of Probability and Expectation Using Computer Simulation Program (Interactive Probability). Master’s Thesis, Korea National University of Education, Cheongju, Republic of Korea, 1999; pp. 9–12.
- Lee, Y.S. A Study of CO₂ Emissions per Vehicle on Roads Using Monte Carlo Simulation. Master’s Thesis, Kyonggi University, Suwon, Republic of Korea, 2012; pp. 17–18.
- Shin, J.Y.; Lee, G.J. Research on application of spatial statistics for exploring spatio-temporal changes in patterns of commercial landuse. *J. Korean Geogr. Soc.* **2007**, *42*, 632–647.
- Manly, B.F.J. *Randomization, Bootstrap and Monte Carlo Methods in Biology*; Chapman and Hall: London, UK, 2006.
- Martin, A. Testing for similarity in area-based spatial patterns: A nonparametric Monte Carlo approach. *Appl. Geogr.* **2009**, *29*, 333–345.
- Seol, Y.J.; Jung, S.B.; Song, K.H.; Chon, K.S.; Rhee, S.M. A evaluation model of AHP results using Monte Carlo Simulation (Depending on the case studies of road and rail). *Korean Soc. Transp.* **2008**, *26*, 195–204.
- Besag, J.; Diggle, P.J. Simple Monte Carlo tests for spatial pattern. *Appl. Stat.* **1977**, *26*, 327–333. [CrossRef]

18. Diggle, P. *Statistical Analysis of Spatial Point Patterns*; Arnold: London, UK, 2003.
19. Loosmore, N.B.; Ford, E.D. Statistical inference using the G or K point pattern spatial statistics. *Ecology* **2006**, *87*, 1925–1931. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.