Resources, Conservation & Recycling 161 (2020) 104884

Contents lists available at ScienceDirect

Resources, Conservation & Recycling

journal homepage: www.elsevier.com/locate/resconrec

Full length article

Development of the WEEE grouping system in South Korea using the hierarchical and non-hierarchical clustering algorithms

Jihwan Park[1,2,*], Keon Vin Park[3], Soohyun Yoo[1], Sang Ok Choi[4], Sung Won Han[3,**]

[1] Research and Development Team, Korea Electronics Recycling Cooperative (KERC), 17 Daehak 4-ro, Suwon, 16226, Republic of Korea
[2] Program in Science & Technology Studies (STS), Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, Republic of Korea
[3] School of Industrial Management Engineering, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, Republic of Korea
[4] Department of Public Administration, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, Republic of Korea

ARTICLE INFO

Keywords:
WEEE
Hierarchical
Cluster analysis
ANOVA
Silhouette, Dunn index

ABSTRACT

South Korea has been operating an extended producer responsibility system (EPR) since 2003 to collect, transport, and dispose of e-waste. Until 2019, the EPR system was operated with a total number of 27 electronic products classified into five categories based on weight and volume, but 23 items will be added in 2020 along with a change to five categories based on the function of the products. In this study that used actual operational data related to the collection, transport, and recycling steps from recycling plants in South Korea, we have analyzed how well the new five-category grouping appropriately reflected actual recycling industrial conditions and have provided optimal classification alternatives. The results showed that clustering accuracy was the best for the classification that used the hierarchical method. In particular, the evaluation index, *silhouettes*, showed the best accuracy with three clusters (0.4155), and the *Dunn index* indicated the best performance with four clusters (0.2333). Based these results, ANOVA tests were implemented, and showed that the three clusters in the relevant models were significantly different with regard to *takt-time, weight, volume*, and *no. of recycling processes* ($p \leq 0.01$) and to both *recycling cost* and *value of material* ($p \leq 0.05$). In contrast, with regard to the grouping suggested by the South Korean government, the overall results of the clustering accuracy using *silhouettes* and *Dunn indices* were –0.2028 and 0.058, respectively. In conclusion, the new grouping suggested by the hierarchical method with four clusters can be utilized as a political decision-making tool.

1. Introduction

In South Korea, a national system for collecting and recycling designated waste of electrical and electronic equipment (WEEE) has been built and operated based on the original Extended Producer Responsibility (EPR) system, which was first introduced in 2003 (Park et al., 2018; Park et al., 2019a). The Ministry of Environment (MOE) has included electrical and electronic equipment (EEE) as a specific target item in the EPR since the test-operation period in 2003 (KEC, 2019a; MOE, 2019). Based on the EPR system, the Eco-Assurance System (Eco-AS) was introduced in 2008 to improve not only the eco-friendly design of EEE products and the recycling rate of WEEE but also to increase the criteria for restricting the use of harmful materials in EEE in the manufacturing step (KEC, 2019b; MOE, 2015). Currently, the Target Management (TM) system, introduced in 2014, is designed to set an annual recycling target quantity from the MOE corresponding to the manufacturers and importers' production and sales volumes based on the previous year. The first target amount sat at 3.9 kg (per capita per year) in 2014, with the amount increasing yearly: 4.5 kg (2015), 4.8 kg (2016), 5.4 kg (2017), 6.0 kg (2018), 6.52 kg (2019), and 7.04 kg (2020) (Park et al.,2020b). Also, at the end of 2018, the MOE announced a long-term recycling target amount of 8.6 kg/cap-yr until 2023.

Since 2014, a total number of 27 EEEs have been included as mandatory target items collected and recycled in four different groups (large-size, mid-size, small-size, and telecommunications) for systematic management under the Eco-AS (Table S1) (Jang, 2010; Manomaivibool and Ho, 2014; Lee and Bae, 2015). However, in order to achieve the long-term recycling target amount, planned for 2023, the South Korean MOE announced that it would expand the number of mandatory EEE items from 27 to 50 items in 2020 (Table 1). The EEE, which had been classified by size in the existing criteria, would be

E-mail addresses: jh.park@k-erc.or.kr (J. Park), swhan@korea.ac.kr (S.W. Han).
* Co-corresponding author: Jihwan Park1, Tel.: +82 31 8014 5441; fax: +82 31 8014 5449
** Co-corresponding author: Sung Won Han3, Tel.: +82 2 3290 3384

https://doi.org/10.1016/j.resconrec.2020.104884
Received 19 December 2019; Received in revised form 7 March 2020; Accepted 13 April 2020

Table 1
Electronic and Electrical Equipment (EEE) Items Subject to Mandatory Collecting and Recycling to be Applied in South Korea from 2020.

Equipment Categories	Specific List of EEE Items	
Temperature Exchange Equipment	Refrigerator (70)	Water Purifier (29)
	Air-conditioner (148)	Dehumidifier (5)
	Vending Machine (6)	
Display Equipment	Television (117)	Computer (Monitor, Laptop) [a] (20)
	Navigation[b] (5)	
Telecommunication Equipment	Copier (20)	Computer (Desktop) [a] (21)
	Printer (23)	Facsimile (15)
	Beam-Projector[b] (5)	Scanner[b] (5)
	Router[b] (5)	Mobile Phone (100)
General Equipment	Oven (16)	Washing Machine (60)
	Microwave (16)	Bidet (34)
	Food-disposal (27)	Dish Dryer (9)
	Heater (17)	Air-cleaner (21)
	Audio (35)	Humidifier (42)
	Rice-cooker (40)	Water Softener (6)
	Iron (48)	Vacuum Cleaner (14)
	Blender (38)	Fan (54)
	Video Player (36)	Toaster[b] (5)
	Kettle[b] (5)	Water Heater[b] (5)
	Frying Pan[b] (5)	Hair Dryer[b] (5)
	Exercise Treadmill[b] (5)	Security Camera[b] (5)
	Food Dryer[b] (5)	Massager (Massage Chair)[b] (5)
	Foot Bath[b] (5)	Sewing Machine[b] (5)
	Videogame Machine[b] (5)	Bread Machine[b] (5)
	Deep fryer[b] (5)	Coffee Maker[b] (5)
	Boiling pot[b] (5)	Dehydrator[b] (5)
Photovoltaic Panel Equipment	Photovoltaic Panel[b] (5)	
Total	51 units[a] (1,197)	

In parentheses, the numbers indicate the samples used in this study's analysis.
[a]Computer was divided into the 'Display' and 'Telecommunication' equipment groups. Thus, a total number of 51 items were categorized.
[b]EEE items that will be newly added as mandatory items in 2020.

newly classified into the functional characteristics of the electronic equipment: temperature exchange equipment, display equipment, telecommunications equipment, general equipment, and solar panel equipment (Park et al., 2019b). The criteria of the EEE classification based on the regulations for the WEEE collecting and recycling has an important simultaneous influence on the collecting/recycling standard cost and the unit cost of the allotted charge announced by the MOE and KERC, respectively. Information about these costs, estimated and announced to be KRW per kilogram (KRW / kg) based on the five different WEEE categories, is the most basic economic data in the area, not only in the implementation of the WEEE resource circulation policy and the EPR system but also as part of the fundamental material for persuading stakeholders allocated in the WEEE collecting, transport, and recycling fields. The collecting/recycling standard and allotted charge costs both consider the actual collecting, transport, and treatment activities along with labor costs and even the valuable price of reproduced resources from the recycling process (Lee et al., 2007; Park et al,.2019b). In other words, although the criteria for the WEEE classification was based on the size of the individual product until 2019, the WEEE will be classified according to the product's functional characteristics starting with 2020, which could lead to a widening gap in volume and weight among items in the same WEEE category. Although not absolute, the volume of a piece of equipment is generally proportional to its weight, and the large volume and heavyweight products require more manpower and more costs during the collecting, transport, and recycling processes. Therefore, the EEE should be classified in consideration of its physical characteristics and recycling-related costs (Park et al., 2019b).

Under the EPR system in South Korea, EEE manufacturers, importers, and sellers have the duty to thoroughly fulfill the responsibilities for WEEE collection and recycling, individually or by joining a Producer Responsibility Organization (PRO), which is the Korea Electronics Recycling Cooperative (KERC). Membership entails paying certain costs to the KERC in the form of "allotted charges," and monies are distributed in the form of "support funds" to the WEEE transporters and recyclers (Lee and Kang, 2016; Park et al., 2018; Park et al.,2019a; Park et al.,2019b). In a short period of time, manufacturers, importers, and sellers each adopted the low-cost case after considering the two options of individually fulfilling their obligations or joining the KERC. Here, the two important costs related to the determination of whether to carry out individual duties or join the KERC are the standard cost of collecting/recycling and the unit cost of the allotted charge (Bahers and Kim, 2018). The standard costs for the collecting/recycling are the legal cost set/announced by the MOE for the collecting and recycling of the WEEE; if the responsible business's collection and recycling quota is not met, a final imposing charge is determined by multiplying the amount not collected/recycled by the standard costs. On the other hand, the allocated charges are calculated by the KERC and are lower than the standard costs. This is because manufacturers, importers, and sellers spend extra money to achieve their obligations individually, such as in constructing infrastructure, but they use a variety of common resources when they join the KERC (Park et al., 2019b). As a result, the KERC focuses on calculating a realistic and reasonable unit cost of allotted charge to attract more manufacturers, importers, and sellers as members of the KERC and conducts an annual analysis of background conditions including collection, transportation, and recycling site surveys. In other words, these cost analysis results from the KERC should be actively utilized in the criteria for the actual product classifications or groupings (MOE, 2015; Rhee, 2016; MOE, 2019).

This study attempted to reclassify products from the five categories of WEEE announced by the South Korean MOE using a cluster analysis approach in order to compare and verify different classifications using the actual WEEE collecting and recycling conditions. In the data acquisition process, unit costs for all the EEE items in the collecting, transport, and recycling processes were investigated and measured at an actual WEEE recycling plant in South Korea. In addition to costs, the number of recycling unit processes (e.g., pre-treatment, shredding, grinding, and sorting), actual takt time (or cycle time), and physical information (e.g., weight and volume) were measured to improve the reliability of the clustering (Lee et al., 2007; Park et al., 2018). Unfortunately, previous studies have rarely examined the rationality of official product classifications through the analysis of the actual recovery, transport, and recycling process conditions of WEEE, in particular. Therefore, the design and implementation of this study were composed and analyzed based on the results of a direct survey of actual recovery, transportation, and recycling plants in South Korea.

Lastly, this is the first study to reinterpret the classification of WEEE in South Korea. Therefore, in this presentation, we begin by organizing the prior research to clearly show the methodology of the relevant clustering analysis and evaluation as well as the basic data on unit process and cost analysis of each product. Then, we organize in detail the methodology for determining the clustering analysis, evaluation methods, and logic based on the collected data. The outcomes obtained from the cluster analysis are presented; based on them, the evaluation indexes were applied to determine the optimal classification system, which is introduced in the results. Information on the new product classification is proposed for national agencies, such as the MOE and the KERC, which can influence policy changes, such as moving products to the military. This is emphasized in the final section that discusses data outcomes and implications for e-waste policies (Figure. 1).

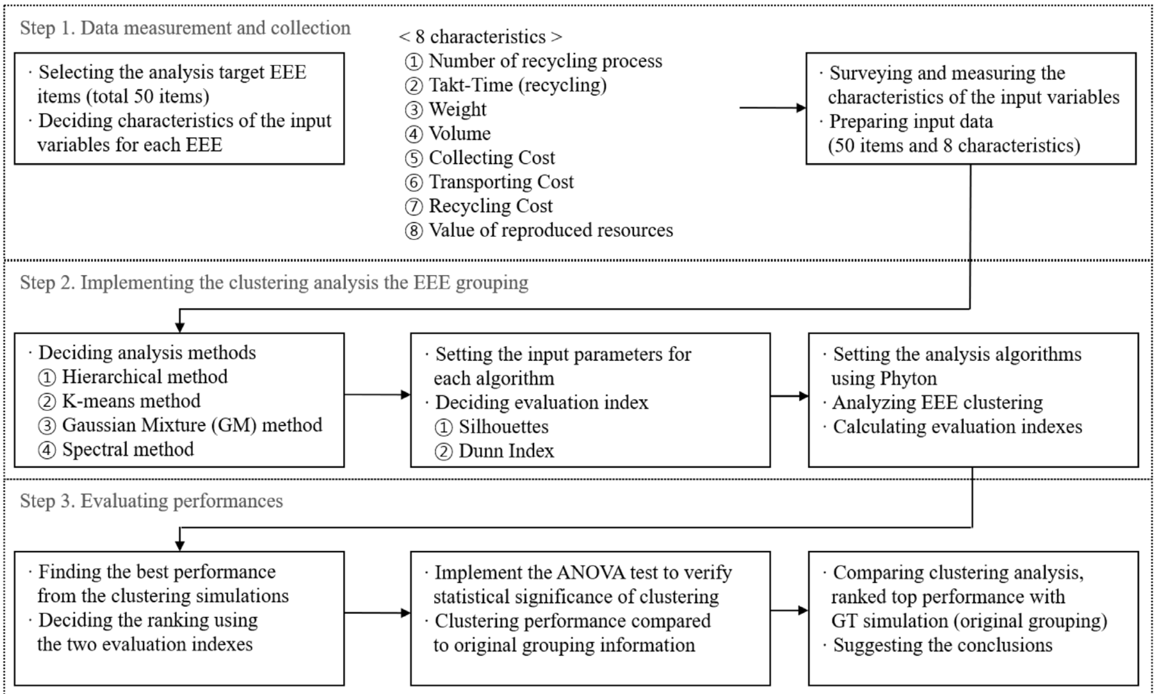

Fig. 1. Schematic flowchart for the cluster analysis in this study.

2. Literature Review

2.1. Cluster analysis

Clustering means finding and defining a number of clusters (or groups) in a dataset. The grouping is done based on numerical similarities or distances (dissimilarities). The inputs required are similarity measures or data from which similarities can be computed [Richard and Dean, 2002]. Cluster analysis has been developed in several different fields with very diverse applications in mind. Therefore, there are a wide range of approaches to cluster analysis and a wider range of methodologies [Hennig, C. et al., 2015]. In general, methods of cluster analysis can be divided into two groups depending on the main concept of the calculation. The first one is hierarchical clustering methods, and the second is non-hierarchical clustering methods. In general, hierarchical clustering methods begin by recognizing each of the initial data elements itself as a separate cluster, and then merges the most similar clusters in an iterative process, gradually creating larger clusters at higher layers. On the other hand, non-hierarchical clustering methods are comprised of various sub-methods, such as the partitioning method, model-based clustering, the graph-based method, and the density-based method. Thus, non-hierarchical clustering methods are much more varied in the algorithms they use to determine a final grouping, and their results vary as well [Wilson et al., 2002; Khan et al., 2019].

At a more detailed level, hierarchical clustering can merge smaller clusters into a larger cluster, or break up a larger cluster into smaller clusters. This is the unique characteristic of hierarchical clustering. When small clusters are merged into larger clusters, it is a bottom-up approach; when large clusters are merged into smaller clusters, it is a top-down approach. In either approach, how related two data elements are (to determine cluster membership) can be determined by a number of different methods, such as Euclidean distance, Manhattan distance, and cosine similarity. In partitional clustering, various partitions are created for the clustering, and the partitions produced are evaluated by several criteria. Partitional clustering is a type of non-hierarchical clustering, and the artificial criteria generated for the partitions affect the placement in the mutually exclusive clusters. Only one cluster set is the result of a typical partitional clustering algorithm. For example, for *k*-means, which is one of the typical non-hierarchical (partitional) clustering methods, users need to have the desired number of cluster inputs, or centroids, which is designated by the number *k*. In other words, the user must pre-determine and enter the number of clusters (*k*) before executing the algorithm, implying that the algorithm starts at the center of the partition with the *k*-value [Fred and Leit, 2000; Wilson et al., 2002; Gülağız and Şahin, 2017].

For non-hierarchical clustering applications, advanced clustering analysis methods have been explored and developed, such as the finite mixture densities and spectral clustering models with their numerous sub-mathematical theories. The basic principle of the finite mixtures (FM) model, one of the model-based clustering methods, is that it is assumed that each observation group providing data suspected to contain clusters comes from a population with a unique probability distribution, and a model suitable for the cluster analysis can be provided. As a commonly utilized example, a Gaussian distribution was applied to the finite mix to measure and analyze the ratio of the body length of 1,000 crabs sampled at Naples (Pearson, 1894). In spectral clustering, points are clustered using the eigenvectors of matrices extracted from the data. Thus, this method is based on graph and matrix theories, and analyzing the Laplacian matrix for obtaining the perfect clustering from the analysis of the eigenvectors (Melnykov and Maitra, 2010).

2.2. Cluster analysis with WEEE

Logistics studies relevant to the topic of EEE or WEEE transport have often used cluster analysis. Guerra et al. (2005) conducted a Rockwell Arena model simulation with clustering approaches to find the optimal reverse logistics paths and the minimum number of vehicles to avoid overlapping vehicles and intervention times at collection centers. Accorsi et al (2012) developed a decision-support system for securing the operational efficiency of reverse logistics for WEEE, based on a mixed-integer linear program and hierarchical clustering methods, and used a clustering approach to identify proper facility locations and match demand and suppliers between the customers and the WEEE collection sites. In South Korea, Lee et al. (2014) implemented a clustering analysis for finding the optimal location for a proposed WEEE recycling plant by considering the information based on the WEEE reverse logistics and actual recycling performances and suggested a new location for the recycling plant with assumptions about the facility capacity and the cost of construction.

Previous studies have long discussed the political and social feature perspectives of WEEE scheme management. Grunow and Gobbi, (2009) developed an optimal decision support system, assigning fair and reasonable individual assignments to all stakeholders, such as producers, transporters, and recyclers, to apply the EU WEEE directives in Denmark. Corsini et al. (2017) implemented a latent class analysis (LCA) with clustering approaches to identify management strategies of EPR based on political conditions, such as policies, supply chain, WEEE collection, and recycling performance, and suggested that nations clarify the principal responsibility to the stakeholders. From the social perspective of WEEE disposal, Lozano et al. (2010) implemented a questionnaire survey and analysis to identify respondents' disposal characteristics when they dumped WEEE. The researchers used a self-organizing map (SOM) clustering method and detected that consumer disposal behaviors were different depending on the specific types of EEE items, with some items (e.g., computers and microwaves) being more commonly properly disposed of than others (e.g., irons, radios, and TVs).

Clustering methods were also used to examine the hazardous risk assessment perspective in studies on the topic of EEE and WEEE. Fujimori et al., (2012) analytically reported heavy metal concentrations (a total of 11 heavy metals) caused by different formal and informal WEEE recycling sites, as measured in the soil and dust in the atmosphere of metro Manila in the Philippines. They also conclusively demonstrated that the concentration of heavy metal from the dust in the formal sector's recycling field was significantly higher than comparative data from other Southeast Asian countries. Also, there are many previous studies that measured and evaluated the concentrations of the heavy metal and organic pollutants from e-waste disposal areas, usually to emphasize the concern for both environmental and occupational hazards (Tang et al., 2010; Pradhan and Kumar, 2014). As seen above, even though clustering techniques have been used in numerous research studies within the WEEE topic, studies related to specific EEE item classification and the management and regulation of WEEE in South Korea have been very rare.

3. Method

3.1. Data acquisition

In the data acquisition step, input variables were selected and investigated for the classification of the WEEE items based on the clustering methods. The input variables that were selected are important factors in the WEEE collection, transport, and recycling process of South Korea and are significantly related to the EEE's physical characteristics, unit costs in various processes, and recycling conditions in the industrial field. In detail, the physical characteristics of the EEE indicated *weight* and *volume* information by actual measurement or specification. The category of the unit costs was divided into three types of costs, depending on the collection *(collection cost)*, transport *(transport cost)*, and recycling (*treatment cost*) processes, respectively. Lastly, the recycling condition indicated *takt-time*, number of recycling processes (*no. of recycling processes)*, and actual *value of reproduced resources* (Table 2). The eight variables just described were also measured and calculated at an actual WEEE recycling plant. The information reflected in these eight variables is unofficially investigated at one time every two years in order to calculate the basic costs of the allotted fund managed by the KERC.

3.1.1. Sampling site and target

Essentially, the fundamental data were investigated, measured, and analyzed from actual WEEE recycling processes of the Metropolitan Electronics Recycling Center (MERC), one of the large-scale WEEE recycling plants in South Korea, directly managed by the KERC. This plant is located in Yongin-si near the capital city of Seoul. MERC first became operational in 2003 and now recycles approximately 22,000 tons of e-waste annually (Park et al., 2018; Park et al., 2019). MERC can provide detailed information on the transport and storage of each EEE product, as well as the actual data on takt-time, unit cost, and the physical characteristics of each treatment (recycling) step. Consequently, we conducted the sampling work to collect information on the eight variables for 50 types of items during the period from December 2017 to December 2018 (one year), targeting a total of 1,197 products. The specific number of samples for each EEE item is shown in Table 1.

3.1.2. Weight and volume

In the data acquisition step, we actually measured volume and investigated weight information for each product. In the case of the weight data, we cited and used 'average weight information' data which are annually reported by the Korea Environment Corporation (KEC) (Table S2). The KEC and KERC measure the average weight per product each year for the statistical management of electronic wastes. The minimum number of samples for measuring weight was in the range from 50 to 200, based on the 'Task Guideline for Recycling and Recycling of Electrical and Electronic Products' in the regulation 'Act on Resource Circulation of Electrical and Electronic Equipment and Vehicles' (KLRI, 2019). However, in the case of mobile phones, the number of samples for the main body and battery was 1,000, and the number for the charger was 300 (TGREP, 2019). The average weight is the total weight of the measured specimens divided by the total quantity of the samples. After calculation, values less than three decimal places in precision were discarded.

For the volume measurement step, we actually measured the volume of each product using a tape measure in MERC. The number of samples used for the volumetric measurements was 10 per product. The most important point of the volume measurement was to establish the criteria for selecting a sample model. In case of refrigerators, we selected samples that had an internal volume from 700 to 800 liters; drum- and general-type washing machines were selected in a 1:1 ratio for measurement. In addition, the volume for televisions (TVs) was calculated by averaging flat-panel display (FPD) and cathode-ray tube (CRT) values; air-conditioners were divided into wall-hangers and stand types by a 1:1 ratio then measured for volume. This ratio will eventually change over time but it was measured actually in recycling center in 2019; Thus, 1:1 ratio perhaps provide insight reflected recycling status in South Korea. Except for refrigerators, air-conditioners, and TVs, other products (almost 47 items) were measured without applying a proportional classification, because they were relatively similar in volume. In common with the weight measurement, the average volume was also finally prepared per each product as an input variable.

3.1.3. Unit cost

For the clustering analysis, three types of unit cost (collecting, transporting, and recycling cost) were selected and collected. Each cost was applied to a standard time principle, which indicated the time in seconds per kilogram of product weight. In other words, the standard time (ST) can be expressed as a unit of sec / kg, and equally applied to the collection, transportation, and recycling processes. Meanwhile, the 'average wage' and 'average working hours' data in specific industries in South Korea, from the Korean Statistical Information Service (KOSIS), were used for the standard cost (SC) calculation as a unit of KRW / sec (KRW is the South Korean won, the official currency). Finally, we multiplied ST and SC by the weight of the product investigated in Section 3.1.2, implying that they consequently produced the unit cost for each collection, transport, and recycle:

$$Unit\ cost = St.\ Time\left(\frac{Sec}{kg}\right) \times St.\ cost\left(\frac{KRW}{sec}\right) \times Average\ weight \tag{1}$$

Table 2
Statistical Summary for the Examined Variables Based on the Original Groupings.

Categories (No. of Items)	Temp. Exchange Equipment (n = 5)	Display Equipment (n = 3)	Telecommunications Equipment (n = 8)	General Equipment (n = 34)	Photovoltaic Panel Equipment (n = 1)	p-values[f]
Takt-Time(min)	16.65 ± 8.91	5.29 ± 3.24	13.43 ± 15.83	11.85 ± 12.84	2.08	0.573000
Weight(kg)	70.27 ± 80.78	13.62 ± 14.46	13.70 ± 26.65	9.10 ± 15.88	18.60	0.000744***
Col Cost(KRW/kg)[a]	68.35 ± 31.34	46.67 ± 34.27	105.5 ± 168.01	53.53 ± 36.54	92.00	0.596000
Trs Cost(KRW/kg)[b]	31.80 ± 15.61	20 ± 10.44	33.25 ± 15.96	37.50 ± 31.60	14.00	0.526000
Rcy Cost(KRW/kg)[c]	268.16 ± 144.72	445.36 ± 455.8	548.17 ± 173.50	635.89 ± 438.22	124.00	0.085700*
Value of Mat (KRW/kg)[d]	513.21 ± 216.24	215.71 ± 124.53	1690.53 ± 3884.33	419.54 ± 196.48	165.00	0.637000
Volume(cm^3)[e]	19.47 ± 1.52	16.75 ± 1.69	15.94 ± 2.60	16.88 ± 2.29	17.48	0.109000
No. of Recycling Processes	7.00 ± 6.40	3.33 ± 0.58	4.75 ± 1.49	5.00 ± 2.31	12.00	0.795000

[a-c]Costs for *collection, transport,* and *recycling,* respectively (100 KRW will be converted to 0.04 USD, criteria in Oct. 2019).
[d]Actual monetary value of resources reproduced by recycling in the recycling plant.
[e]Units of the volume variable were transformed to log.
[f]All variables were tested by ANOVA.
*p-value ≤ 0.10, **p-value ≤ 0.05, ***p-value ≤ 0.01.

In other words, the unit costs for the collection, transportation, and recycling processes depend on the applicable scope of each unit process and the weight of the products. The ranges for the three processes are as follows. First, the range of the collection costs covered the cost of shipping from the consumer's house to the distribution (storage) center and the storage fee there (assuming each product was disposed). Since various collection channels could not be surveyed nationwide, the collection point most similar to the distance average was selected from about five collection paths that returned to the Yongin-si storage site. After investigating the distance, fuel costs, and the actual average wage data of the driver from the KOSIS, the ST and SC were finally calculated. Second, the transport cost was calculated by determining the move from a warehouse or storage site to a recycling center, including the round-trip transport and the loading and/or landing fee. In this study, transportation costs were converted to ST- and SC-like collection costs by selecting the best warehouse or storage site for the distance average out of about 30 routes transported to the MERC from various locations. Lastly, the treatment cost, which means the cost required for the WEEE recycling process operation, was defined as the recycling cost in this study. The recycling cost consisted of the direct and indirect labor costs for the workers, performance incentives, overhead, operating costs for equipment/facilities, and waste disposal costs. Information on recycling cost was calculated by measuring the ST and SC based on the actual MERC operation data. Similar to the above collection and shipping costs, the average salary was based on the industry average data from the KOSIS.

3.1.4. Takt-time and value of resources

The takt-time aggregated all the time spent on the final recycling process from the WEEE incoming process to the recyclable-resources-producing process in the recycling center. The takt-time of the recycling process was investigated accordingly and recognized as a very important variable, because it reflected the actual recycling conditions. The takt-time data were investigated for each of the 50 products, and the time for each step was measured using actual operational data and video from cameras, which were added together. There were 10 processes included in the takt-time: 1)WEEE unloading in the RC, 2) classification and temporary storage in placement, 3) movement to the plant, 4) movement to the work space, 5) (recycling process) pre-treatment, 6) (recycling process) conveying to the next step, 7) treatment (i.e., shredding, dismantling), 8) sorting (mechanically or manually), 9) resources packaging, and 10) resources loading to truck (recyclable-resources-producing process in the recycling center).

The value of resources reproduced is very important data for carefully considering the recycling conditions in South Korea, because not only can the expected revenue be determined when the resources reproduced by WEEE recycling are traded in the market, but it can also provide background knowledge for supporting decision-making, for example, in helping to decide which recycling types to apply to increase expected revenue by selling recyclable resources. The potential value of resources reproduced for each WEEE item was calculated in the following order. First, each sampled WEEE mentioned in Section 3.1.1 was directly dismantled and broken down to determine the proportion and weight (unit : kg) of specific components (e.g., ferrous metals, plastic, copper, aluminum, etc.). Second, the actual transaction or market price (unit : KRW / kg) of recyclable components reproduced by MERC was investigated. Finally, the potential value of one product was calculated by multiplying the major component and weight of each product by the market price.

3.1.5. Number of recycling processes

The number of recycling processes (*no. of recycling processes*) variable was quantified based on the number of unit processes in the real recycling stage. In fact, all WEEE are recycled in mechanical, semi-mechanical, and manual (disassembly) types of processes. Electronic products that are to be mechanically recycled in a recycling plant involve numerous recycling unit processes. In contrast, electronic products that are to be disassembled manually have to undergo a simple process for recycling in a plant. For this research, we investigated the 50 products being recycled in MERC by examining the number of unit processes being applied in the recycling process, regardless of the mechanical, semi-mechanical and manual types for each product, and used the results.

3.2. Clustering algorithms

A total of eight variables and their information were investigated and collected for 50 electronic items (with five groups) through the data acquisition process as mentioned in Section 3.1. As a result, we implemented a clustering algorithm, examined the appropriateness of the initial five groups and a number of newly proposed groups, and studied the characteristics of the products grouped together in the new results. This study attempted to derive the best solution by applying various clustering principles and comparing the results. To do this, we used hierarchical, partitioning, model-based, and graph-based clustering algorithms. In other words, the final clustering techniques were hierarchical with Ward's algorithm (hierarchical method), k-means (partitioning method), Gaussian mixture model (model-based method), and spectral clustering (graph-based method).

3.2.1. Hierarchical clustering algorithm

Ward considered hierarchical clustering procedures based on minimizing the 'loss of information' from joining two groups [Ward, 1963]. This method is usually implemented with a loss of information taken to be an increase in an error sum of squares criterion, *ESS*. Thus, *ESS* is defined as follows:

$$ESS = \sum_{k=1}^{K} \sum_{x_i \in C_k} \sum_{j=1}^{n} (\boldsymbol{x}_{ij} - \mu_{kj})^2 \tag{2}$$

where K is the number of clusters, j is the sample index$(1, 2, \ldots, n)$,

C_k is a k-th cluster,

x_i indicates elements in C_k,

x_{ij} denotes the i-th item in the j-th cluster C_j, and

μ_{kj} is a mean vector in C_k.

Initially, each cluster consists of a single item. At each step in the analysis, the union of every possible pair of clusters is considered, and the two clusters whose combination results in the smallest increase in *ESS* (a minimum loss of information) are joined. When all the clusters are combined in a single group of N items, the clustering analysis is done. The results of Ward's method can be displayed as a dendrogram. The vertical axis gives the values of *ESS* at which the merges occur [Richard and Dean, 2002].

3.2.2. K-mean clustering algorithm

Assume we have a dataset $\{x_1, x_2, \ldots, x_n\}$, which are n observations consisting of a random p-dimensional vector. Our goal is to cluster the data into K clusters, where the value of K is pre-determined. We might think of a cluster as comprising a group of data points whose inter-point distances are small compared with the distances to points outside of the cluster. We can formalize this logic by first introducing a set of p-dimensional vectors $\mu_k = \{\mu_{k1}, \mu_{k2}, \ldots, \mu_{kp}\}$, where $k = 1, \ldots, K$ in which μ_k is a mean vector of the kth cluster. Our objective is to find an assignment of data points to clusters as well as a set of vectors μ_k that the within-cluster sum of squares is minimized. For each data point x_n, we introduce a corresponding set of binary variables $r_{nk} \in \{0, 1\}$ where $k = 1, \ldots, K$ describing which of the K clusters the data point x_n is assigned to, so that if data point x_n is assigned to cluster k then $r_{nk} = 1$ and $r_{nj} = 0$ for $j \neq k$. We can then define an objective function as follows:

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} x_n - \mu_k, \tag{3}$$

$$\mu_k = \{\mu_{k1}, \mu_{k2}, \ldots, \mu_{kp}\}, \ \ where\ k = 1, \ldots, K,$$

$$r_{nk} \in \{0, 1\}, \ \ where\ k = 1, \ldots, K, \ \ n = 1, \ldots N, \ \ and$$

$$x_n = \{x_{n1}, x_{n2}, \ldots, x_{np}\}, \ \ where\ n = 1, \ldots, N.$$

In Equation (3), our goal is to find values for $\{r_{nk}\}$ and $\{\mu_k\}$ to minimize J. A solution can be found using an iterative procedure involving two steps corresponding to optimizations with respect to $\{r_{nk}\}$ and $\{\mu_k\}$ [Bishop, 2006].

3.2.3. Gaussian mixture algorithm

The Gaussian mixture is widely used as one of the clustering algorithms. The form of the Gaussian mixture distribution is formulated by the parameters π, μ, and Σ, where we have used the notation $\pi = \{\pi_1, \pi_2, \ldots \pi_k\}$, $\mu = \{\mu_1, \mu_2, \ldots \mu_k\}$, and $\Sigma = \{\Sigma_1, \Sigma_2, \ldots \Sigma_k\}$. Moreover, π (pi), μ (mu), and Σ (SIGMA) are sets of respectively, the relative weights, means, and covariance matrices of the superimposed multidimensional Gaussian distributions that are expected to be in the data. The aim is to find the parameter sets $\{\pi, \mu, \Sigma\}$ that best fit the data, which is achieved by maximizing a log-likelihood function. The log of the likelihood function is given by [Bishop, 2006]:

$$\ln p(X|\pi, \mu, \Sigma) = \sum_{n=1}^{N} \ln\left\{\sum_{k=1}^{K} \pi_k \mathrm{N}(x_n|\mu_k, \Sigma_k)\right\} \tag{4}$$

where $X = \{x_1, x_2 \ldots x_N\}$.

In Equation (4), the maximum likelihood solution for the parameters does not have a closed-form solution. One approach to maximizing the log-likelihood function is to use iterative numerical optimization techniques.

3.2.4. Spectral clustering algorithm

Spectral clustering is one of the graph-based clustering algorithms. It uses a weighted graph to cluster data points. A weighted graph is specified by $\mathbb{G} = (\mathbb{V}, \mathbb{E}, \mathrm{W})$ where $\mathbb{V}$ is the set of all nodes; $\mathbb{E}$ is the set of edges connecting the nodes; and W is an affinity matrix with weights characterizing how likely two nodes are to belong in the same group. We applied a Gaussian similarity function (x) with $\sigma = 1$ to define the weights (as it is one of the most widely used similarity functions) [Von, 2007]. Let $[n]$ denote the set of integers between 1 and n: $[n] = \{1, 2, \ldots, n\}$. Let $\mathbb{V} = [N]$ denote the set of all elements (data points) to be grouped. To cluster N points into K groups is to decompose $\mathbb{V}$ into K disjoint sets, in other words, $\mathbb{V} = \cup_{l=1}^{K} \mathbb{V}_l\ and\ \mathbb{V}_k \cap \mathbb{V}_l = \varnothing,\ \forall\ k \neq l$. We denote this K-way partitioning by $\Gamma_V^K = \{\mathbb{V}_1, \mathbb{V}_2, \ldots, \mathbb{V}_K\}$. Let $\mathbb{A}, \mathbb{B} \subset \mathbb{V}$. We define links$(\mathbb{A}, \mathbb{B})$ to be the total weighted connections from $\mathbb{A}$ to $\mathbb{B}$:

$$\mathrm{links}(\mathbb{A}, \mathbb{B}) = \sum_{i \in \mathbb{A}, j \in \mathbb{B}} W(i, j)\ , \tag{5}$$

where $W(i, j) = \exp(\frac{-d(x_i, x_j)^2}{2\sigma^2})$ is the Gaussian similarity function, and $d(x_i, x_j)$ is the Euclidean distance between two data elements.

The degree of a set is simply the total links to all the nodes:

$$\mathrm{degree}(\mathbb{A}) = \mathrm{links}(\mathbb{A}, \mathbb{V}). \tag{6}$$

Using the degree as a normalization term, we define:

$$\mathrm{linkratio}(\mathbb{A}, \mathbb{B}) = \frac{\mathrm{links}(\mathbb{A}, \mathbb{B})}{\mathrm{degree}(\mathbb{A})}. \tag{7}$$

The linkratio$(\mathbb{A}, \mathbb{B})$ in Equation (7) means the proportion of the links with $\mathbb{B}$ among those $\mathbb{A}$ has. Two special link ratios are of interest. One is linkratio$(\mathbb{A}, \mathbb{A})$,which measures how many links stay within $\mathbb{A}$ itself. The other is its complement linkratio$(\mathbb{A}, \mathbb{V}\setminus\mathbb{A})$ which measures how many links escape from $\mathbb{A}$. A good clustering desires both tight connections within partitions and loose connections between partitions. This objective can be accomplished by maximizing linkratio$(\mathbb{A}, \mathbb{A})$ and minimizing linkratio$(\mathbb{A}, \mathbb{V}\setminus\mathbb{A})$ [Yu, 2003].

3.3. Clustering evaluation indices

After developing and optimizing the cluster algorithms, the adequacy of the analysis methods were compared using clustering validity indices, which determine how useful the results of the cluster creation are. Cluster relevance indicators comprehensively consider the distance between clusters, the diameter of clusters, and the dispersion of clusters. In other words, cluster adequacy is assessed based on inter-group variance and intra-group variance. In this study, two indicators were used: the *Dunn index* and *silhouettes*.

3.3.1. Dunn index

The *Dunn index* is an indicator that denotes the minimum value of the distance between clusters and the maximum value of the distance between elements in a cluster. In other words, the longer the distance between the clusters and the smaller the clusters, the better the clustering. In this case, the *Dunn index* will become larger with better clustering. The index definition is given by the following:

$$\mathrm{D} = \frac{\min\limits_{i \neq j}\{d_c(C_i, C_j)\}}{\max\limits_{1 \le l \le k}\{\Delta(C_l)\}} \tag{8}$$

where $d_c(C_i, C_j) = \min\limits_{x \in c_i,\ y \in c_j}\{d(x, y)\}$ or the distance between two clusters,

$\Delta(C_l) = \max_{x,\, y \in c_l} \{d(x, y)\}$ or the diameter of C_l,

$d(x, y)$ is the Euclidean distance between two data elements, and

k is the number of clusters.

In Equation (8), if a dataset contains well-separated clusters, the distances among the clusters are usually large and the diameters of the clusters themselves are expected to be small. Therefore, a larger value means a better cluster configuration (Kovács et al., 2005)

3.3.2. Silhouettes

Silhouettes are used to evaluate how well clustering results are clearly separated. In order to construct *silhouettes*, we only need two things, the partition and the collection of all proximities between objects. For each object i, we will introduce a certain value $s(i)$. Take any object i in the dataset and denote by A the cluster to which it has been assigned. Also, $a(i)$ is the average Euclidean distance of *object* i to all other objects of A, $b(i)$ is the minimum of the average Euclidean distance between i and the objects in other clusters to which the object i does not belong. Then, $s(i)$ is obtained by combining $a(i)$ and $b(i)$ as follows in Equation (9). Finally, the *silhouettes* index takes the average of $s(i)$ for all the samples to evaluate the clustering result, like in Equation (10) (Rousseeuw, 1987). The *silhouettes* index ranges from -1 to 1, where a high value indicates better clustering accuracy.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i),\, b(i)\}} \tag{9}$$

$$Silhouette = \frac{1}{\mathrm{n}} \sum_{i=1}^{n} s(i) \tag{10}$$

4. Results

4.1. Statistics summary

A total of 1,197 samples from the 50 items were collected and experimentally examined for the cluster analysis. The quantity of products collected and analyzed for each of the 50 products was the same as shown in Table 1. For the 23 items that will be newly added to the mandatory list in 2020, the experimental analysis was conducted by securing at least five samples for each product (due to the difficulty in securing a larger number of samples). For 27 products, which were mandatory substitute items as of 2019, a larger quantity had been obtained and analyzed, ranging from at least six (vending machines) to 148 (air conditioners, including indoor and outdoor appliances) samples. With the initial information on the five groups (defined by the MOE), the eight variables were investigated for each of the 50 products, and the information was integrated into the group and summarized in Table 2. The temperature-exchange group was relatively higher than the other groups in terms of takt-time (16.65 ± 8.91 min), weight (70.27 ± 80.78 kg), and volume (19.47 ± 1.52 cm^3) for the log-transformed data; the collection cost (105.5 ± 168.01 KRW / kg) and values of the materials (1690.53 ± 3884.33 KRW / kg) reproduced in the telecommunications group were also relatively more expensive than the other groups. The cost of transportation (37.50 ± 31.60 KRW / kg) and recycling (635.89 ± 438.22 KRW / kg) in the general equipment group was relatively higher than that of the other groups. Regarding the number of recycling processes, the temperature-exchange group had more recycling processes (7.00 ± 6.40) than the other groups.

Through ANOVA testing, a total of eight variables were tested for significant differences among the groups. The results showed that the *weight* was significantly different among the groups at a p-value less than the 0.01 level, and the *recycling cost* was also found to differ significantly among the groups at a p-value less than the 0.10 level. From the statistics summary (as shown in the Table 2), we have concluded that the original groups of the EEE products presented by the MOE can only be considered appropriately distinct classifications in terms of *weight* ($p \le 0.01$) and *recycling cost* ($p \le 0.10$). However, it is hard to conclude that appropriate product line classifications have been made, because there were no other significant differences among the groups in terms of *takt-time, collection cost, transport cost, value of the material reproduced, volume,* and *no. of recycling processes* with p-values at a significance level of no more than 0.10.

According to the correlation analysis using the eight variables, the *volume* of each product had significant correlations with all the variables except two; *no. of recycling processes* and *treatment cost* were not significantly correlated with the *volume* of each product with p-values at a significance level of more than 0.10 (Table S3). This implies that the volume of the product was proportional to the *takt-time* and *weight* but inversely proportional to the costs (collection and treatment) and value of the reproduced material. In addition, *weight* and *takt-time*, and the *value of the material reproduced* and *collection costs* were each positively correlated, respectively ($p \le 0.01$). Finally, in terms of recycling difficulty, takt-time was relative to the weight and volume of the product. In other words, since the recycling takt-time of relatively heavy and bulky products is increased, these products can be inferred as having recycling intractability or complexity.

4.2. Clustering analysis

Hierarchical (hierarchical-based), k-means (partitioning-based), spectral (graph-based), and Gaussian mixture (model-based) clustering methods were empirically evaluated, and the validity of each resulting clustering model was also evaluated using *silhouettes* and the *Dunn index*. In the cluster analysis, the number of clusters was varied from two to six for each method. In the case of the hierarchical method, the number of clusters was varied from two to six based on the error sum of squares, which can be used as evidence for the evaluation of the clustering. As shown in Table 3, the four clustering methods were evaluated by *silhouettes* (in the upper half of the table) and by the *Dunn index* (in the lower half of the table).

The evaluation results of the clustering using the *silhouettes* were as follows. The results of the hierarchical method had the best performance with four clusters, and the clustering validities in terms of *silhouette* values were as follows (from highest to lowest): 4 clusters (0.4155), 3 (0.3790), 2 (0.3603), 6 (0.2479), and 5 (0.2422). In the case of the k-means, the clustering performance in terms of *silhouette* values was as follows (from highest to lowest): 4 clusters (0.3601), 3 (0.3514), 2 (0.3331), 5 (0.2305), and 6 (0.2147). In the case of the spectral clustering and Gaussian mixture, the results showed the best *silhouette* values with three clusters. Spectral clustering was evaluated in terms of clustering performance in decreasing *silhouette* value order: 3 clusters (0.3943), 2(0.3775), 4 (0.2668), 6 (0.1952), and 5 (0.1713). Gaussian mixture performed the best with a *silhouette* value of 0.2690 with three clusters. The clustering accuracy in other cases depended on the number of clusters; in decreasing *silhouette* value order, they were: 6 clusters (0.2624), 5 (0.2538), 2 (0.2007), and 4 (0.1972). The best techniques according to *silhouette* values were (in decreasing order) hierarchical, spectral clustering, k-means, and Gaussian mixture; four clusters resulted in the best *silhouette* value for the best technique, the hierarchical method.

On the other hand, the overall results of the cluster assessment using the Dunn index differed from the results based on the *silhouette* value. With the Dunn index, the optimal number of clusters was three or four in the hierarchical method. The highest *Dunn index* was 0.2333, when three or four clusters were chosen using hierarchical. The cluster effectiveness of the hierarchical solutions as measured by the Dunn-index were as follows (from highest to lowest): 3 or 4 clusters (0.2333), 2 (0.1675), 6 (0.1548), and 5 (0.1380). In the case of k-means, the best performance with *Dunn index* was also shown with three or four clusters, and the specific results were as follows (from highest to lowest): 3 or 4 clusters (0.1709), 6 (0.1409), 5 (0.1272), and 2 (0.1227). The spectral clustering results for the *Dunn index* were (from highest to

Table 3
Clustering Analysis Method Simulation Results According to Number of Communities (Set Parameter).

Evaluation with *Silhouette*	Hierarchical Clustering[a] (Ward Method)	*k*-means Clustering	Spectral Clustering	Gaussian Mixture
2	0.3603 (ESS = 2.2)[a]	0.3331	0.3775	0.2007
3	0.3790 (ESS = 2)	0.3514	0.3943	0.2690
4	<u>0.4155</u> (ESS = 1.5)	0.3601	0.2668	0.1972
5	0.2422 (ESS = 1.3)	0.2305	0.1713	0.2538
6	0.2479 (ESS = 1.16)	0.2147	0.1952	0.2624
Evaluation with *Dunn Index*	**Hierarchical Clustering[a] (Ward Method)**	***k* -means Clustering**	**Spectral Clustering**	**Gaussian Mixture**
2	0.1675 (ESS = 2.2)[a]	0.1227	0.1675	0.0978
3	<u>0.2333</u> (ESS = 2)	0.1709	0.1675	0.1055
4	<u>0.2333</u> (ESS = 1.5)	0.1709	0.0935	0.1250
5	0.1380 (ESS = 1.3)	0.1272	0.0890	0.1360
6	0.1548 (ESS = 1.16)	0.1409	0.1464	0.1360

The underlining indicates the highest values of the *Silhouettes* and *Dunn Index*, respectively.
[a]The *ESS* (error sum of squares) was applied to the hierarchical clustering method only.

lowest): 2 or 3 clusters (0.1675), 6 (0.1464), 4 (0.0935), and 5 (0.0890). For Gaussian mixture, the best performance was a *Dunn index* of 0.1360 with five or six clusters, but this was somewhat lower than the hierarchical, *k*-means, and spectral methods; for the other numbers of clusters, the *Dunn index* value order decreased, for 4 clusters (0.1250), 3 (0.1055), and 2 (0.0978). According to the Dunn index, the hierarchical clustering results for three or four clusters were equally good, whereas the silhouettes indicated that the best result contained four clusters. Although the optimal number of clusters with the best performance differed slightly, there was not a dramatic difference between the *silhouettes* and *Dunn index* results because the four-cluster solution of the hierarchical method had the best performance in both.

To sum up the above results, in the case of the evaluation by *silhouettes*, the clustering performance was the highest with 0.4155, when 50 products were grouped into four clusters by the hierarchical method. Also, the clustering performance when separating the EEE products into three clusters by the spectral method followed next with 0.3943. Meanwhile, the assessment using *silhouettes* showed the classification results of the electronic equipment were the difference for each clustering method. On the other hand, as a result of the clustering evaluation using the *Dunn index*, when the 50 products were classified into three or four clusters using the hierarchical method, the model showed the highest performance of 0.2333. Subsequently, using the *k*-means method, the 50 products showed good model performance when also classified into three or four clusters (*Dunn index* = 0.1709).

4.3. Grouping with four clusters

Using the hierarchical method, 50 types of electronic products were confirmed to be best grouped when the *ESS* value was 1.5 in Table 3. Accordingly, the classification characteristic was analyzed assuming the number of groups was four. The newly classified clusters based on the hierarchical method, when the *ESS* was 1.5 (4 clusters), are shown in Table 4. Group 1 includes a total of 3 EEE products, Group 2 includes 7 products, Group 3 includes 40 products, and Group 4 includes 1 product, respectively (Table 4). First, the major characteristics of the mobile phone with details are shown in Table 5. Mobile phones have a relatively short takt-time (0.13 min), very light weight (0.24 kg), and small volume (10.67 cm^3) when compared with other WEEE items (the last two were log transformed as previously noted). Also, the number of recycling processes was six stages; these include, for example, manually detaching the battery and PCB boards from the main body of the phone, the main body is put into a mechanical shredder to prevent personal information leakage, and the crushed matter goes to sorting (magnetic and eddy-current) processes to separate ferrous and non-ferrous materials, and finally, recyclable resources were released to other plant. In

Table 4
New Grouping Results Based on Hierarchical Clustering Analysis with Highest *Silhouette* and *Dunn index* values (*ESS* = 1.5).

Equipment Categories	Specific List of EEE Items	
Group 1 (*n* = 3)	Refrigerator (70)	Washing Machine (60)
	Photovoltaic Panel[b] (5)	
Group 2 (*n* = 7)	Vending Machine (6)	Air-conditioner (148)
	Copier (20)	Washing Machine (60)
	Oven (16)	Exercise Treadmill [b] (5)
	Sewing Machine[b] (5)	
Group 3 (*n* = 40)	Water Purifier (29)	Dehumidifier (5)
	Television (117)	Computer (Monitor, Laptop)[a] (20)
	Computer (Desktop)[a] (21)	Navigation[b] (5)
	Printer (23)	Facsimile (15)
	Scanner[b] (5)	Beam-projector [b] (5)
	Router[b] (5)	Food-disposal (27)
	Microwave (16)	Air-cleaner (21)
	Bidet (34)	Audio (35)
	Heater (17)	Water softener (6)
	Rice-cooker (40)	Iron (48)
	Humidifier (42)	Blender (38)
	Fan (54)	Video Player (36)
	Vacuum Cleaner (14)	Water Heater [b] (5)
	Kettle[b] (5)	Hair Dryer [b] (5)
	Frying Pan[b] (5)	Food Dryer[b] (5)
	Security Camera[b] (5)	Massager (Massage Chair)[b] (5)
	Foot Bath[b] (5)	Bread Machine[b] (5)
	Videogame Machine[b] (5)	Coffee Maker[b] (5)
	Deep Fryer[b] (5)	Dehydrator[b] (5)
	Boiling Pot[b] (5)	Toaster[b] (5)
Group 4 (*n* = 1)	Mobile Phone (100)	
Total	51 units[a]	

[a]Computer was divided into the 'Display' (monitor) and 'Telecommunications' (main body) equipment groups, thus, a total of 51 items were analyzed.
[b]EEE items will be new mandatory items in 2020.

terms of cost, transport (28 KRW / kg) and recycling costs (580 KRW / kg) were comparable to products of other groups, but the collection cost was about 10 times more expensive than other products, and the value (11,300 KRW / kg) of resources reproduced was about 30 times more valuable than other products (Table 5). There is a very important point of interest from these cluster results. That being that even though *mobile phone* was a single item, it was segregated into one group (cluster) with no other members based on its recycling characteristics. This result indicated that the cluster differences among the two- or three-cluster

Table 5
Clustering Analysis Results of the Hierarchical Method, the Best Performance with Four Grouping (*ESS* = 1.5).

Category	Cluster 1 (n = 3)	Cluster 2 (n = 7)	Cluster 3 (n = 40)	Cluster 4 (n = 1)	p-value[f]
Takt-Time(min)	6.21 ± 3.66	38.40 ± 13.40	8.10 ± 5.21	0.13	0.001230***
Weight(kg)	54.47 ± 33.03	62.75 ± 68.84	5.67 ± 5.89	0.24	7.94e-06***
Col Cost(KRW)[a]	73.00 ± 17.06	52.25 ± 31.80	53.45 ± 36.34	515.00	0.119000
Trans. Cost(KRW)[b]	21.67 ± 7.51	31.71 ± 21.65	36.48 ± 29.38	28.00	0.415000
Recycle Cost(KRW)[c]	133.33 ± 8.14	369.91 ± 196.05	630.93 ± 413.49	580.00	0.014400**
Value of Mat. (KRW)[d]	301.33 ± 119.51	379.04 ± 218.55	407.73 ± 196.54	11300.00	0.024700**
Volume (cm^3)[e]	19.48 ± 1.80	19.31 ± 1.36	16.56 ± 1.98	10.67	3.5e-05***
No. of recycling processes	15.00 ± 3.00	4.29 ± 1.50	4.60 ± 1.53	6.00	3.34e-06***

[a-c]Cost for *collection, transport*, and *recycling*, respectively (100 KRW will be converted to 0.04 USD, criteria in Oct. 2019).
[d]Actual monetary value of resources reproduced by recycling in the recycling plant.
[e]Units of the volume variable were transformed to log.
[f]All variables were tested by ANOVA.
**p-value ≤ 0.10, **p-value ≤ 0.05, ***p-value ≤ 0.01.*

solutions and the four-cluster solutions produced different *silhouettes* and *Dunn index* values depending on whether or not mobile phones were included, and it can be inferred that a single product can significantly affect the clustering classification.

Table 5 shows the clustering results when *ESS* = 1.5 through the hierarchical method. Cluster 1 had relatively lower costs than the other clusters, transportation costs (21.67 ± 7.51 KRW / kg) and recycling costs (133.33 ± 8.14 KRW / kg), except for collection costs (73.00 ± 17.06 KRW / kg). In addition, Cluster 1 had less value (301.33 ± 119.51 KRW / kg) of the materials reproduced by the recycling process than the other clusters, and the number of recycling processes (15.00 ± 3.00) was relatively larger than that of the other groups, implying that this WEEE could be recycled by using both (or mixed) mechanical facilities and manual dismantling methods. In Cluster 2, takt-time (6.21 ± 3.66 min) was relatively longer than the other three clusters, and the number of recycling processes (4.29 ± 1.50) was relatively smaller, indicating that electronic products in Cluster 2 could be recycled and manually dismantled in an actual recycling plant. In Cluster 3, the data for physical characteristics, such as weight (5.67 ± 5.89 kg) and volume (16.56 ± 1.98 cm^3) with log transformed variables, were relatively lower than the other clusters. In terms of the number of recycling processes in Cluster 3 (4.60 ± 1.53), all the items (n = 40) were usually recycled manually but some electronic products could be recycled by mechanical facilities in a recycling plant due to value of the *no. of recycling processes* being relatively higher than those of Cluster 2 (4.29 ± 1.50) that were categorized by the recycling type of a manual dismantle feature. Lastly, with the inter-clusters' ANOVA test, group differences for the eight variables were analyzed. According to the results of the ANOVA test, the variables *takt-time, weight, volume*, and *no. of recycling processes* had significant differences among the groups (clusters) at a significance level of less than 0.01, and the two variables, *recycling cost* and *value of reproduced materials*, were also significantly different among the groups (clusters) at a significance level of less than 0.05. Visualizations of the clustering results using the three clusters are shown in Figure 2 with respect to the variables between *volume* and *takt-time* and *volume* and *transport-cost*.

Cluster 1 consisted of the products that had the largest volumes and heaviest weights compared to the other two clusters. The average takt-time was relatively short due to the adoption of the mechanical recycling type, and the resulting deviation was small. Refrigerators and washing machines are currently being collected and transported with the same products by a palletizing method in South Korea. In particular, refrigerators and washing machines have low collection and transport costs relative to other products due to their large amount of reverse logistics. Currently, there is only one factory that mechanically processes solar panels in South Korea for recycling, so there is no basis for comparison. However, the actual results show that they are recovered and shipped as a single product, so they are less costly in terms of collection and transport cost than many other products. One of the characteristics for a number of the seven items in Cluster 2 is that they have relatively large volumes and heavy weights compared to Cluster 3. In other words, numerous quantities can be collected and transported as a single item compared to those in Cluster 1, thus making them relatively cheaper to collect, transport, and recycle. However, they had recyclable properties using manual work, and the takt-time spent on recycling was the longest among the three groups. In other words, the value of the materials generated by recycling the WEEE was offset by the lower working efficiency among the three groups due to their longer takt-time. After all, it can be judged that it is a product line that needs to make efforts to develop the most intensive management and dedicated automation processes to achieve the national recycling quota in South Korea. With regard to Cluster 3, this group includes a number of electronics with a small size, low weight, and a small volume. In other words, this cluster, based on the current mandatory items and groups (27 items with 5 groups), is comprised of a mix of mid-size and small-size telecommunication products, but they are composed mainly of products with light weight and small volume. These products naturally tend to increase collection, transport, and recycling costs, because smaller volumes and more types decrease the chance that large quantities of e-waste will be collected as a single product. The more difficult it is to collect; the more products will eventually be mixed (Table 5). As additional collection activities are needed, not only collection and transport costs will be increased, but also recycling costs are also increased due to high labor costs as people have to manually disassemble the products in the recycling plant.

By summing up the study results, the overall meaningful conclusions of this study are displayed in Table 6. First, the original clusters proposed by the South Korean MOE were evaluated using *silhouettes* and the *Dunn index* for cluster integrity, with –0.2028 and 0.0585 respectively. Given the range of values in *silhouettes* and the *Dunn index*, the above results were assessed to have very low clustering accuracy. However, when creating a four-cluster model using the hierarchical method on the same dataset, the *silhouettes* and *Dunn index* improved to 0.4155 and 0.2333, implying that these models demonstrated an empirically high clustering accuracy (Table 6).

5. Conclusion and Limitations

5.1. Conclusion

In this study, we collected actual and operational data related to WEEE recycling in situ from collection to recycling and conducted a cluster analysis on a total of 50 WEEE products that are included in the collection and recycling obligations. The purpose of the cluster analysis was to contribute to the e-waste recycling scheme and performance

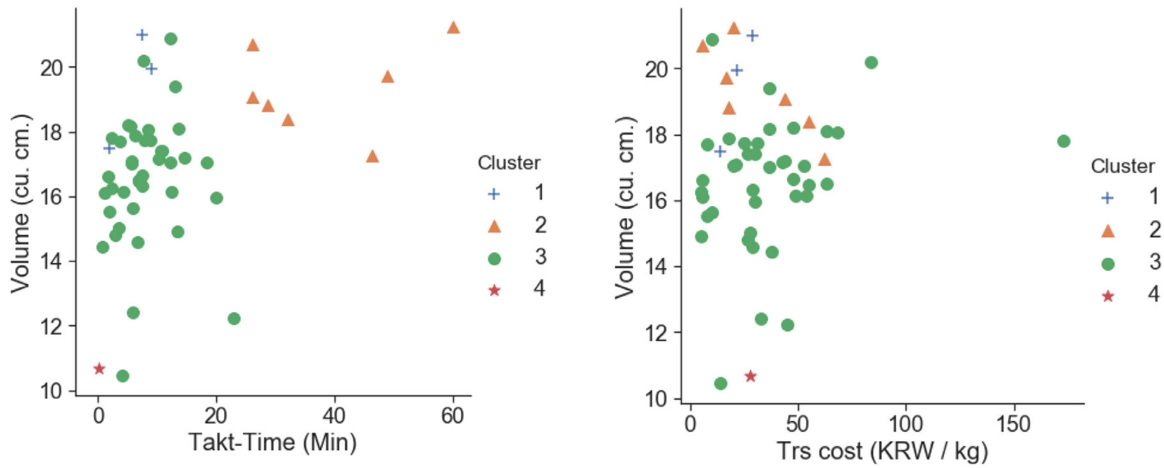

Fig. 2. Results of the clustering using hierarchical methods (*ESS* = *1.5*); clustering plots with respect to *volume* and *takt-time* (left side) and to *volume* and *transport cost* (right side).

Table 6
Performance Comparison Between Original Grouping with Clustering Suggested by Using *Silhouettes* and *Dunn Index*.

Equipment Categories	Original Group (No. of Clusters = 5)	Hierarchical Clustering (No. of Clusters = 4)
Silhouette[a]	-0.2028	0.4155
Dunn Index[b]	0.0585	0.2333

[a]has the range from -1 to 1.
[b]has the range from 0 to infinity.

improvement by proposing a verification of the five groupings presented by the South Korean MOE and comparing it with a statistically significant grouping. A total of eight variables were selected as input variables considering the physical characteristics (*weight* and *volume*) of the EEE items, three different types of unit costs (*collection, transport*, and *recycling costs*) from the collection to the recycling stage, and actual recycling conditions (*recycling type, takt-time*, and *value of materials reproduced*). A total of 1,197 products were sampled and analyzed to build a dataset on these variables. Meanwhile, for the clustering analysis, four different cluster models were built based on the fundamental principles of clustering, hierarchical (hierarchical method), partitioning (*k*-means), graph (spectral method), and model-based (Gaussian mixture), and the clustering results were evaluated using *silhouettes* and the *Dunn index*, two strong indicators for assessing clustering accuracy.

Using the four different cluster models, the results of the cluster analysis showed that the highest *silhouette* value was 0.4155 with four clusters (*ESS* = 1.5) in hierarchical method. Meanwhile, the highest *Dunn index* was 0.2333, two cases, when the numbers of clusters were three (*ESS* = 2) and four (*ESS* = 1.5) in the hierarchical method, respectively. A total of 50 electronic products were classified by the hierarchical method; the hierarchical method showed the best performance in both evaluation cases, using the *silhouette* and *Dunn index*. One of the important points obtained through the cluster analysis was as follows. The results of the hierarchical method with the best accuracy (*ESS* = 1.5) indicated that the highest *silhouettes* and *Dunn index* were when the number of clusters was four with mobile phones forming a one-group as the only member of one cluster. In other words, the different results between the three- and four-cluster solutions in the hierarchical method depending on the practical difference of whether or not the single item of a mobile phone was separated into its own unique cluster. Meanwhile, based on the eight variables with the actual data collected, the five groups proposed by the Korean MOE were analyzed, and the results showed that the values of *silhouettes* and *Dunn index* were very low (-0.2028 and 0.0585, respectively). when the number of clusters was four (*ESS* = 1.5), for the hierarchical method.

We need to consider ways to apply the important results of this study in the operation of e-waste policies and systems in South Korea. South Korea's e-waste recycling system is based on EPR and Eco-AS. Beginning in 2020, the number of items eligible for mandatory recycling will be increased to 50. Importantly, South Korea's collection and recycling performance are managed and reported based on a ‘product group.’ In addition, the standard-cost that is imposed if recycling obligations are not fulfilled will be also applied differently by product group. This means that in the actual field of the WEEE collection and recycling industry, the government's product grouping and its criteria can be a critical determinant of the business. After all, this research is very meaningful in that it was designed and conducted based on actual data gathered on the electronics recovery and recycling industry in South Korea. In particular, if certain products are required to move or change grouping due to their characteristics, at that time, the results of this study may be used as a basis for such decisions.

Here are some relevant conclusions:

- Overall information on the five groups proposed by the Korean MOE is not reflective of the overall recycling industry and product-specific recycling conditions as have been detailed in this study. However, it is strictly classified based on the functions of the product (temperature exchange, display, telecommunications, general, and PV panel), therefore, the Environment Ministry can clarify the rationale and justification for this group classification.
- Photovoltaic panels can be managed with other products when reflecting on the recycling industry and product specific conditions. Even if recycling is automated (mechanical) and the collection scheme or system were different, you can determine that they can be managed along with products, such as refrigerators and washing machines.
- It is recommended that mobile phones be managed separately (as a separate group). Mobile phones are worth more than their weight and volume indicate, so they are expensive items. If mobile phones are managed in the same group with other products, they can cause inequalities in the areas of recovery and recycling costs (the basic unit of imposed fines if obligations are not met).
- This study was mainly measured in actual recycling plants in South Korea. All of the data used in the study were actual data. In other words, this research data and the corresponding results can be used as an important basis for calculating unit cost of 'allotted funds' and 'supporting costs in collection and recycling' for agencies, such as the Producer Responsibility Organization (PRO).

5.2. Limitations

In the data acquisition stage, a total of 1,197 samples were collected and analyzed in this study during the period of about a year. At present, 27 products, which are mandatory for recycling, were sampled in numbers from 6 to 148, but only five samples were collected and analyzed for 23 new products to be added for 2020. As a result, the number of samples for all 50 products could not be standardized. In a further study, additional samples for the 23 new items will need to be collected and analyzed. Another limitation, when measuring product volume, two different types of products, such as TVs (CRT and FPD) and air conditioners (hanger and stand), were analyzed as averages based on their actual incidence (1:1). It may change (1:2 or 1:3) over time. As a result, continuous observations and periodic studies of changes in consumption and recycling status are needed. In the data acquisition stage, is uncertainty about the variable 'number of recycling processes.' In this study, the number of recycling processes for 50 products was investigated and used as a variable. This number is accurate but lacks a high degree of precision, as we cannot yet fully define the number of steps in the recycling process due to the wide arrange of categories and individual products within categories. To make up for these weaknesses and improve the quality of the data, a survey will be conducted on about 60 domestic recycling centers or recycling companies through further investigation of the standard recycling process for each item. As these additional improvements are completed, we expect that further analysis will continue to be added to the literature base expanded by the current study.

We did not provide all information related to the eight variables for the 50 electronic products. However, the investigated data for the eight variables were provided in terms of each group (as presented in Tables 2 and 5 to inform the readers). The analyses focused on the group data, and we have already discussed the limitations involved in the product sampling; it was meant to be representative, not comprehensive. Further research might be able to present better methods and more refined information in this area. In the results section, there is one particular limitation in the ANOVA test presentation. We did not provide results of post-hoc comparison tests following the significant main effects. We believed the provided tables and its information allowed us to observe the inter-cluster differences in the context of the investigated variables. We were also primarily concerned that the clustering was creating distinct categories and that the distinctions were relevant with respect to the context of the recycling industry.

We used the *silhouettes* and *Dunn index* as evaluation indices to assess the clustering accuracy. However, as *silhouettes* and the *Dunn index* were introduced in the method portion of Section 3, those indexes cannot be directly compared to each other, since not only their computation methods but also the ranges of the indices are different. Therefore, they should be used as independent evaluation indices for determining the clustering accuracy in a complementary manner. For this reason, the optimal classification was chosen with the number of four clusters from the hierarchical method using the highest indexes from both.

All authors agree with the CRediT author statement

- The corresponding author is responsible for ensuring that the descriptions are accurate and agreed by all authors.
- The role(s) of all authors should be listed, using the relevant above categories.
- Authors may have contributed in multiple roles.
- CRediT in no way changes the journal's criteria to qualify for authorship.

CRediT author statement

Jihwan Park : Conceptualization, Writing- Original draft, Writing - Review & Editing.
Sung Won Han : Project administration, Supervision.
Keon Vin Park : Software, Formal analysis, Writing- Original draft.
Soohyun Yoo : Investigation, Data Curation
Sang Ok Choi : Project administration, Resources.

Declaration of interests

• The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

•The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Acknowledgements

"This research was supported by the Korea University Graduate School Junior Fellow Research Grant."

References

Accorsi, R., Manzini, R., Bortolini, M., 2012. A hierarchical procedure for storage allocation and assignment within an order-picking system. A case study. International Journal of Logistics Research and Applications 15 (6), 351–364. https://doi.org/10.1080/13675567.2012.742877.
Bahers, J., Kim, J., 2018. Regional approach of waste electrical and electronic equipment (WEEE) management in France. Resources, Conserv. Recycl. 129, 45–55. https://doi.org/10.1016/j.resconrec.2017.10.016.
Bishop, Christopher M., 2006. Pattern recognition and machine learning. New York, NY: Springer.
Corsini, F., Rizzi, F., Frey, M., 2017. Extended producer responsibility: The impact of organizational dimensions on WEEE collection from households. Waste Manag 59, 23–29. https://doi.org/10.1016/j.wasman.2016.10.046.
Fujimori, T., Takigami, H., Agusa, T., Eguchi, A., Bekki, K., Yoshida, A., Terazono, A., Ballesteros, F.C., 2012. Impact of metals in surface matrices from formal and informal electronic-waste recycling around metro Manila, the Philippines, and intra-Asian comparison. J. Hazard. Mater. 221–222, 139–146. https://doi.org/10.1016/j.jhazmat.2012.04.019.
Grunow, M., Gobbi, C., 2009. Designing the reverse network for WEEE in Denmark. CIRP Annals 58 (1), 391–394. https://doi.org/10.1016/j.cirp.2009.03.036.
Guerra, L., Murino, T., Romano, E., 2005. Reverse logistics for electrical and electronic equipment: A modular simulation model. 307–312.
Gülağız, F.K., Şahin, S., 2017. Comparison of hierarchical and non-hierarchical clustering algorithms. International Journal of Computer Engineering and Information Technology 9 (1), 6–14.
Jang, Y., 2010. Waste electrical and electronic equipment (WEEE) management in Korea: Generation, collection, and recycling systems12, 283–294. https://doi.org/10.1007/s10163-010-0298-5.
Khan, K.S., Yin, Y., Jamalipour, A., 2019. On the application of agglomerative hierarchical clustering for cache-assisted D2D Networks. In: 16th IEEE Annu. Consum. Commun. Netw. Conf, pp. 1–6.
Kovács, F., Legány, C., Babos, A., 2005. In: Cluster validity measurement techniques. 6th International Symposium of Hungarian Researchers on Computational Intelligence, pp. 35.
Lee, Chan Hee, Bae, Wookeun, 2015. A comparative study on the waste electrical and electronic equipment (WEEE) recycling policies of the European Union and the Republic of Korea. J. Kor Envir Polic and Admin. 23 (1), 43–68. https://doi.org/10.15301/jepa.2015.23.1.43.
Lee, J.C., Song, H.T., Yoo, J.M., 2007. Present status of the recycling of waste electrical and electronic equipment in Korea. Resour. Conserv. Recycl. 50, 380–397. https://doi.org/10.1016/j.resconrec.2007.01.010.
Lee, S.K., Roh, J-W, Cho, Y.B., 2014. A study on the optimal method for recycling the waste electronics' reverse logistics. Int J. Comme. Inf. Rev. 16 (4), 171–190.
Lee, W.Y.J., Kang, S.P.J., 2016. Gasification applicability study of polyurethane solid refuse fuel fabricated from electric waste by measuring syngas and nitrogenous pollutant gases. J. Mater. Cycles Waste Manag. 18, 509–516. https://doi.org/10.1007/s10163-016-0512-1.
Lozano, S., Esparza, J., Adenso-Diaz, B., Garcia, J.M., 2010. Clustering Spanish households e-waste disposal behavior using self- organizing feature maps. In: 2010 IEEE International Conference on Industrial Engineering and Engineering Management, pp. 2328–2332. https://doi.org/10.1109/IEEM.2010.5674177.
Manomaivibool, P., Ho, J., 2014. Two decades, three WEEE systems: How far did EPR evolve in Korea's resource circulation policy ? Resour. Conserv. Recycl 83, 202–212. https://doi.org/10.1016/j.resconrec.2013.10.011.
Melnykov, V., Maitra, R., 2010. Finite mixture models and model-based clustering. Stat. Surv. 4, 80–116. https://doi.org/10.1214/09-SS053.
Park, J., Ahn, C., Lee, K., Choi, W., Teak, H., Ok, S., 2019a. Analysis on public perception, user-satisfaction, and publicity for WEEE collecting system in South Korea: A case study for door-to-door service. Resources, Conserv. Recycl. 144, 90–99. https://doi.org/10.1016/j.resconrec.2019.01.018.
Park, J., Jung, I., Choi, W., Ok, S., Won, S., 2019b. Greenhouse gas emission offsetting by

refrigerant recovery from WEEE: A case study on a WEEE recycling plant in Korea. Resources, Conserv. Recycl. 142, 167–176. https://doi.org/10.1016/j.resconrec.2018.12.003.

Park, J., Jung, I., Lee, K., Kim, M., Hwang, J., Choi, W., 2018. Case study in Korea of manufacturing SRF for polyurethanes recycling in e-wastes. J. Mater. Cycles Waste Manag. 20, 1950–1960. https://doi.org/10.1007/s10163-018-0718-5.

Pradhan, J.K., Kumar, S., 2014. Informal e-waste recycling: Environmental risk assessment of heavy metal contamination in Mandoli industrial area, Delhi, India. Environ. Sci. Pollut. Res 21, 7913–7928. https://doi.org/10.1007/s11356-014-2713-2.

Richard, A.J., Dean, W.W., 2002. Applied multivariate statistical analysis. Prenticee Hall, London, pp. 265.

Rhee, S., 2016. Beneficial use practice of e-wastes in Republic of Korea. Procedia Environ. Sci. 31, 707–714. https://doi.org/10.1016/j.proenv.2016.02.055.

Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 20, 53–65.

Tang, X., Shen, C., Chen, L., Xiao, X., Wu, J., Khan, M.I., Dou, C., Chen, Y., 2010. Inorganic and organic pollution in agricultural soil from an emerging e-waste recycling town in Taizhou area, China 895–906. Journal of Soils and Sediments 10, 895–906. https://doi.org/10.1007/s11368-010-0252-0.

Ward, Joe H., 1963. Hierarchical grouping to optimize an objective function. J. of the Americ. stat. Assoc. 58 (301), 236–244.

Wilson, H.G., Boots, B., Millward, A.A., 2002. A comparison of hierarchical and partitional clustering techniques for multispectral image classification. IEEE Int. Geosci. Remote Sens. Symp. 3, 1624–1626. https://doi.org/10.1109/IGARSS.2002.1026201.