ABSTRACT  The reduction in government support and the rapid growth in the number of nonprofit organizations have made them face fierce competition for charitable donations. Identifying valuable donors and developing effective marketing strategies can contribute to online donation platforms. This study explored online donors’ characteristics in Taiwan through the identification of different donor segments using a refined clustering algorithm. Furthermore, the marketing strategies based on the salient features of each segment are offered to retain donors and maximize their monetary donations. A real dataset derived from 14,029 donation records contributed by 7,432 donors during the years of 2016-2018 on an online donation platform were collected. A refined cluster analysis based on an improved particle swarm optimization algorithm was applied according to RFM (Recency, Frequency, and Monetary) values and donors’ socio-demographic variables (e.g., Sex, Age, and Education). The results offered four segments of online donors in Taiwan. “Passive donors” were found to be the largest segment (38%), followed by “female active donors” (24%), “potential donors” (21%), and “male loyal donors” (17%). Most donors on the platform were female, highly educated, and aged between 30 and 40. The men’s single donation amount was higher than women’s; however, the women’s total donations were higher than men’s. We contributed the donor segmentation process with a refined clustering technique, which combines RFM and socio-demographic variables as criteria to compensate for the shortcomings of previous studies that only focused on RFM.

INDEX TERMS  Online donor segmentation, donation marketing strategies, socio-demographic factors, RFM.

I. INTRODUCTION  Nonprofit organizations (NPOs) have faced fierce competition for charitable donations because of the decrease in governmental support and the rapid growth in the number of NPOs [1]. Therefore, achieving a bigger portion of monetary donation is the main goal of NPOs. In addition to the traditional channel for donations, online donations are a new channel due to the popularity of the Internet. Independent donors have tended to be the major contributors in Taiwan. At present, an average of more than 200,000 people make donations each year, with a total annual amount of up to NT$45.8 billion. The amount of regular donations is an average of NT$656.4 per person per month [2]. The re-donate rate reduced from 29% to 25% from 2017 to 2018, and the proportion of new and old donors in the two years also dropped from 58 (new donors) to 42 (old donors) in 2017, and 64 (new donors) to 36 (old donors) in 2018 [2]. The evidence indicates...
that the donation amount has declined, and donation behavior has not been consistent in recent years. Accordingly, identifying valuable donors and developing effective marketing strategies can contribute to online donation platforms. Donor segmentation is a marketing tactic applied to effectively improve fundraising for charitable organizations [3]. Such techniques allow for effective resource allocation and donor prioritization [4].

The common problem in segmentation is to select one or more segmentation variables which are suitable for use in the context of donation [5]. Rupp et al. [6] showed that charitable donors have been segmented by a range of demographic, psychographic and behavioral factors, as suggested by several studies related to customer segmentation [7], [8]. In terms of behavior, RFM (Recency, Frequency, Monetary) is a powerful and well-known analytical method for segmentation due to its reflection of customer behavior (Durango-Cohen et al., 2013). Previous studies used customer transaction data to extract RFM and segment them for most marketing goals. Numerous studies have utilized RFM for consumer segmentation in various fields, such as banking, retailing, alumni reunion fundraising, etc. [9]–[11]. Nonprofit organizations, in particular, have relied on RFM analysis to target donors among those who have been the source of contributions in the past [6], [12]. Accordingly, the RFM values were selected as the criteria for segmentation of charitable donors in this study.

Other than donors’ behavioral factors, numerous studies have attempted to segment charitable donors by utilizing their demographic factors as the criteria [6], [13]. The sociodemographic variables such as age, gender, and education achievements are the critical indicators of donors’ charity giving [14]. For example, men with a higher educational level donate more than women with a high education level [15]. Previous studies explored the charity giving either according to the factor of donors’ behavior patterns or individual factors; however, there is little research combining RFM values and socio-demographic variables as the criteria for donor segmentation. Srnka et al. [16] revealed that sociodemographic characteristics combined with behavioral variables can be an efficient criteria for potential donor segmentation. Furthermore, Sarvari et al. [17] demonstrated that the performance of customer segmentation can be enhanced by combining RFM and demographic factors. A donors’ profile can be understood from individual factors and behavior patterns. Accordingly, donor segmentation was implemented based on the criteria of socio-demographic variables and RFM values in this study. Both factors addressed in the study have generated interest in taking a more systematic approach to understanding online donors.

Clustering can group samples into specific clusters according to the pattern of data without labeled classes because the cluster pattern could have potential signification within a specific problem. Samples have a high degree of similar patterns in a cluster but have a high degree of dissimilar patterns to samples of other clusters. Therefore, clustering reflects data structures to facilitate the recognition of potential patterns present in a data set. One of the techniques useful for segmenting donors is the clustering which has been widely used in various studies [3], [18], [19]. The $K$-means algorithm has been widely used in customer and donor segmentation [9]; however, it has poor implementation in related studies concerning clustering analysis [20], [21]. Unlike previous studies which adopted the $K$-means algorithm as the clustering technique, this study applied an improved particle swarm optimization (PSO) with the aim of increasing the accuracy of the clustering. In addition, the donors’ sociodemographic variables (sex, age, educational achievement) and behavioral variables (recency, frequency, monetary) were selected as the criteria to increase the fitness of the segmentation. Numerous clustering approaches have been proposed, including $K$-means [22], PSO, and Gauss chaotic mapping (GCM) PSO (GCMPSO) [20], [21]. These approaches have been successfully applied to many practical problems [23], [24], and have been successfully used to develop the RFM model to identify customer behavior [25], [26].

The purpose of this study is to apply the clustering technique for online donor segmentation based on sociodemographic variables and RFM (Fig. 1). We used the GCMPSO clustering algorithm [20], [21] to determine the optimal clustering solution, which was selected from 2–7 clusters according to the shortest sum of the intracluster distances of all clusters. The error rate and minimum percentage of the sample in $K$ clusters were used to investigate online donors’ characteristics and donation behavior in Taiwan through the identification of various online donor segments. The identified segments can be used as a reference to develop appropriate communication, recruitment, and retention strategies for online donation platforms. Marketing strategies were formulated according to the features of each segment to maintain the relationship between the donors and the platform and to enhance the retention of the donors. Real donation data on online donation platforms in Taiwan were used for segmentation in this study. The profiles of online donors in Taiwan were investigated, and marketing strategies targeting segments with various sociodemographic features that can effectively aid in charity fundraising are suggested.

The remainder of this paper is organized as follows. The relevant approach is summarized in Section II, where we define the Socio-demographic variables and RFM, and design of segmentation analysis and marketing strategies. Case studies and result analyses are provided in Section III. In Section IV, we discuss the advantages of a method combining the RFM values and socio-demographic variables to segment donors using a PSO algorithm for clustering and marketing implications. Finally, we conclude this paper.

II. METHODS

A. SOCIO-DEMOGRAPHIC VARIABLES

Potential donor segmentation could be grouped by the extrinsic measures or the intrinsic variables [14]. Extrinsic variables reflect the donor’s characteristics, while the intrinsic variables reflect the donor’s motivation and personality. Additionally, this study utilized demographic variables (sex, age, educational achievement) and behavioral variables (recency, frequency, monetary) as the criteria for potential donor segmentation. The relevant approach is summarized in Section II, where we define the Socio-demographic variables and RFM, and design of segmentation analysis and marketing strategies. Case studies and result analyses are provided in Section III. In Section IV, we discuss the advantages of a method combining the RFM values and socio-demographic variables to segment donors using a PSO algorithm for clustering and marketing implications. Finally, we conclude this paper.
measures represent demographic and socioeconomic profiles of the charity donors, and intrinsic determinants address the psychographic variables for donors. In addition, a literature review of donor segmentation showed that socio-demographic, psychographic and behavioral variables are usually selected as the criteria of segmentation [6], [27]. Among the socio-demographic criteria, age, gender and educational attainment were the most commonly used variables [6].

Educational attainment is an important predictor of donation behavior [18], [28]. Previous studies showed that individuals with higher educational attainment and who are older tend to make more charitable donations than those with lower educational attainment and who are younger [29]. This behavior could be explained by the fact that donors with higher educational attainment tend to understand others’ needs and thus have greater willingness to help [18], [30]. The individuals with higher educational attainment have more financial capital, which in turn provides more resources to donate [31]. With regard to age, it has been proposed that older adults donate more because of lifecycle effects [13], [18]. Finally, the relationship between gender and charitable donation is significant. For instance, women demonstrate a greater likelihood to give than men [13], [14], [29], but men give higher amounts on average than women [32], [33]. Given the evidence, the potential roles of each of the aforementioned socio-demographic variables were selected as criteria for clustering in this study.

**B. RFM**

In addition to demographic factors, donor behavior is often considered a criterion for segmentation [3], [34]. The donors’ behavior could be described based on the RFM model. The RFM value is a preferred method of segmentation [35], and considered as an easy and effective technique for defining customer segmentation [36]. The RFM analysis proposed by Hughes [37] is a method that differentiates important customers from transaction data according to three attributes. The first dimension is recency, which indicates the length of time since the start of a transaction. The second dimension is frequency, which indicates how frequently a customer purchases products during a particular period. Finally, monetary value measures the amount of money that the customer spends during a certain period of time. According to the literature [38], the higher the value of recency and frequency, the more likely the corresponding customers are to produce new trade with enterprises. Moreover, the greater the monetary value, the more likely the corresponding customers are to buy products or services with enterprises again. In many practices, RFM analysis is carried out by using clustering methods to categorize customers into several segments based on the RFM variables, subsequently linking the segments to the respective average CLV (customer lifetime value). This framework provides deeper insights into customer profitability if we can derive useful associations between customer characteristics and segment membership [39].

**C. DESIGN OF SEGMENTATION ANALYSIS AND MARKETING STRATEGIES**

All steps of the design of segmentation analysis and marketing strategies are illustrated in Fig. 1. First, the goal of donor segmentation was identified. Next, the data collection and cleaning were performed. The donation records from 2016 to 2018 were selected, and missing data and outliers in the donation records were eliminated. The donation records were aggregated into the RFM values from the same individual. Further, the clustering was processed based on the GCMPSO algorithm, and the clusters were grouped as segments according to the values of RFM in each cluster. The saliences of socio-demographic values were considered to label the donor segments. Finally, the marketing strategies were provided based on the features of each segment.
1) PSO

PSO [40] is a swarm-based optimization approach, where a swarm consists of \( N \) particles, denoted as \( S = (p_1, p_2, \ldots, p_N) \). A \( D \)-dimensional feasible solution is regarded as a particle consisting of a set of \( D \)-dimensional decision variables \( x \), denoted as \( p_i = \{1, 2, \ldots, x_{iD}\} \) where \( x \in (X_{\text{min}}, X_{\text{max}})^D \). The search space is the feasible region of the problem defined by the set of all feasible solutions. For optimized solutions, the vector of a particle in the swarm is adjusted in the search space during iterations. The vector of a particle \( p_i \) is adjusted by a velocity (denoted as \( v_i = \{1, 2, \ldots, v_{iD}\} \), where \( v_i \) refers to one velocity and two vectors, including the old velocity of \( p_i \), the best previously visited position of \( p_i \) (denoted as \( pbest_i = \{pbest_{i1}, pbest_{i2}, \ldots, pbest_{iD}\} \)), and the global best position among the swarm (denoted as \( gbest = \{gbest_1, gbest_2, \ldots, gbest_D\} \)). The equation of adjusted vector can be formulated as:

\[
\begin{align*}
v_{i,d}^{t+1} & = w \times v_{i,d}^t + c_1 \times r_1 \times (pbest_{i,d} - x_{i,d}^t) + c_2 \times r_2 \times (gbest_d - x_{i,d}^t) \\
x_{i,d}^{t+1} & = x_{i,d}^t + v_{i,d}^{t+1}
\end{align*}
\]

(1)

where \( x_{i,d}^{t+1} \) and \( v_{i,d}^t \) are the adjusted \( d^{th} \) element and the current \( d^{th} \) element in the position of the \( i^{th} \) particle, respectively. \( v_{i,d}^{t+1} \) and \( v_{i,d}^t \) are the adjusted \( d^{th} \) element and the current \( d^{th} \) element in the velocity vector of the \( i^{th} \) particle, respectively. \( r_1 \) and \( r_2 \) are random numbers between \((0, 1)\). \( c_1 \) and \( c_2 \) are the constants to influence the velocity of a particle in one iteration. \( t \) is the current number of iterations and \( w \) is the inertia weight that influences \( v_{i,d}^t \) of the \( i^{th} \) particle. An inertia weight decreases linearly from 0.9 to 0.4 throughout the search process (Eq. 3) and is used to effectively balance exploration and exploitation [41].

\[
w = (0.9 - 0.4) \times \frac{\text{iteration}_{\text{max}} - t}{\text{iteration}_{\text{max}}} + 0.4
\]

(3)

where \( t \) is the current number of iterations; \( \text{iteration}_{\text{max}} \) is the maximum number of iterations. The PSO pseudocode is illustrated in Algorithm 1.

2) GAUSS CHAOTIC MAPPING PSO

In PSO, \( r_1 \) influences the particle updating related with \( pbest \), and \( r_2 \) influences the particle updating related with \( gbest \). GCM can generate chaotic sequences based on the chaotic behavior mapping. In Gauss chaotic map PSO (GCMPSO), two sequences \( g_1 \) and \( g_2 \) generated by the GCM are used in place of \( r_1 \) and \( r_2 \) to improve the balance between the global exploration and the local search ability. \( g_1 \) and \( g_2 \) are formulated by

\[
g_{i,d}^{t+1} = \left\{ \begin{array}{ll}
0, & g_{i,d}^{t-1} = 0 \\
\text{Frac} \left( \frac{1}{g_{i,d}^{t-1}} \right), & g_{i,d}^{t-1} > 0
\end{array} \right.
\]

(4)

where \( \text{Frac}(\cdot) \) returns the fractional portion of a scalar, \( t \) is the current number of iterations, and \( g_{i,d}^{t-1} \) is a sequence \( g \) in the \( d^{th} \) element of the \( i^{th} \) particle in the current number of iterations. Thus, Eq. (1) can be improved using Eq. (4) as follows:

\[
v_{i,d}^{t+1} = w \times v_{i,d}^t + c_1 \times g_{1,i,d}^t \times (pbest_{i,d} - x_{i,d}^t) + c_2 \times g_{2,i,d}^t \times (gbest_d - x_{i,d}^t)
\]

(5)

3) GAUSS CHAOTIC MAPPING PSO ALGORITHM FOR CLUSTERING

The pseudocode of GCMPSO for clustering is illustrated in Algorithm 2. The encoding, fitness evaluation, and \( pbest \) and \( gbest \) updating of particles need to be defined for the clustering problem.
Step 1. Initialize particles. The particle is defined by $K$ center positions for $K$ clusters, where $K$ is the total number of clusters. Let $F$ be the total number of columns in the dataset. The $D$ dimension of a particle is $K \times F$. The $F$ intervals in $D$ represent a center point. A vector $\{x_{i1}, x_{i2}, \ldots, x_{iD}\}$ and velocities $\{v_{i1}, v_{i2}, \ldots, v_{iD}\}$ of a particle are randomly generated within the search space.

Step 2. Evaluate fitness. Fitness is used to determine the value of particles among the swarm. Equation 6 is the fitness evaluation which is the sum of the intra-cluster distances of all clusters. The lower sum of distances is the lower error rate.

$$\text{fitness} = \sum D(x_j, z_j),$$

where $i = 1, 2, \ldots, K$, $j = 1, 2, \ldots, L$. In Eq. 6, $K$ is the total number of clusters; $L$ is the total number of samples in the dataset. $Z_i$ is the position of the $j^{th}$ cluster center and $X_j$ is the position of the $j^{th}$ sample. The intra-cluster distances are calculated by Euclidean distance (Eq. 7). A matrix $X_j \in (C_1, C_2, \ldots, C_K)$, where $C_i$ is the $i^{th}$ cluster centroid vector among $K$ clusters, is used to calculate the distance as the length between $X_j$ and $Z_i$ (Eq. 8). The sample in the dataset is assigned into a cluster according to the shortest distance.

$$D(x_j, z_i) = \sqrt{\sum_{d=1}^{D} (x_{jd} - z_{id})^2}$$

$$Z_i = \frac{1}{n_i} \sum_{\forall x_j \in C_i} x_j$$

Step 3. Update $pbest$ and $gbest$. According to the fitness value, a lower sum of distances indicates a lower error rate. A particle compares the current fitness and the fitness of $pbest$. Both the fitness and position of $pbest$ are replaced by the fitness and position of $p_i$ when the fitness of $p_i$ is a lower value than the fitness of $pbest$; otherwise, the fitness and position of $pbest$ are not changed. Both the fitness and position of $pbest$ are replaced by the fitness and position of $pbest$; when the fitness of $pbest$ is a lower value than the fitness of $gbest$; otherwise, the fitness and position of $gbest$ are not changed.

Step 4. Parameter settings. The maximum iteration is set to 100 and the total number of particles is set to 50. Both $c_1$ and $c_2$ are set to 2. $V_{max}$ is equal to $(X_{max} - X_{min})$ and $V_{min}$ is equal to $-(X_{max} - X_{min})$ [40].

III. CASE STUDIES

A. DATASET

A real case study with actual donation records is adopted to demonstrate the results based on the RFM and socio-demographic criteria by clustering for donors’ segmentation. The data were collected from a famous online donation platform named NPO channel (https://www.npochannel.net/) in Taiwan. NPO channel is mainly used to assist NPOs in raising charitable funds, such as announcing fund-raising projects, and collecting monetary donations through the platform. The NPO channel is one of the biggest online donation platforms in Taiwan; therefore, the donors on the platform sufficiently represent the profiles of Taiwan online donors. The donation records with missing values and outliers were eliminated first. The NPO channel was developed in June 2012. From 2012–2015, external powers, such as advertising, social news, and charitable events, drove the charitable donation of the channel. From 2016 to 2018, the charitable donation of the platform was driven completely by internal powers (e.g., attractiveness of charitable projects, donors’ motivates) without external powers. Therefore, the donation records during 2016 to 2018 presenting truly donation behavior were selected for analysis in the study. Among all the collected donation data, a final set of 14,029 donation records contributed by 7,432 donors from 2016 to 2018 were analyzed. The record for each donor includes the donor’s identifier, sex, age, education, recency, frequency, and monetary value (Fig. 2). Recency is the number of days between the last day of donation and the end day of 2018. Frequency refers to the total number of donations from 2016 to 2018. Monetary value refers to the total amount given by each donor from 2016 to 2018. Monetary value refers to the total amount given by each donor from 2016 to 2018. First, the RFM values for each donor were calculated separately from the data. Second, clustering was executed based on six attributes (e.g., Sex, Age, Education, recency, frequency, monetary). Finally, the researcher and expert worked together to group and label the donor segmentation based on the RFM and socio-demographic values of each segment.

B. OPTIMAL CLUSTERING SELECTION

In this study, we evaluated $K = 2–20$ clusters in the GCMPSO algorithm to select the optimal clustering solution. The error rate and minimum percentage of the sample in $K$ clusters were selected to determine the optimal clustering solution. Fig. 3 represents the error rate, maximum percentage of the sample, and minimum percentage of the sample for $K = 2–20$. The GCMPSO algorithm exhibits a lower error rate for $K = 5$ and 9–20 clusters than for $K = 2, 3, 4, 6, 7$ clusters. A high minimum percentage of the sample was obtained for $K = 2–6$ clusters. The results revealed that a large number of clusters is not necessary in the analysis because the number of samples in a cluster is insufficient to obtain a solution. Moreover, a ten-fold cross-validating approach was used to validate the optimal number of clusters. The 90% dataset was used to select the optimal number of clusters, and this process was repeated ten times to obtain ten results of optimal clustering. The results indicated that $K = 5$ was the best clustering solution in terms of both the error rate and minimum percentage of the sample.

C. COMPARISON OF K-MEANS AND GCMPSO ALGORITHMS

The $K$-means algorithm was used to generate the $K = 2–20$ clusters (Fig. 4). For a minimum percentage of the sample in the $K$ clusters of the $K$-means algorithm, the results obtained when $K = 2–6$ were superior to those obtained when $K = 7–20$. The optimal clustering solution was
D. DONOR SEGMENTATION

The process of donor segmentation is as follows. First, clusters were found by using the clustering technique according to the RFM and socio-demographic criteria. Figure 5 shows the RFM bubble chart of the five clusters, and Table 1 shows the basic features of each cluster.

Recency across the five clusters ranges from 266.52 to 823.14, and is ranked as cluster1, cluster3, cluster0, cluster4, and finally cluster2. The donation frequency across all clusters ranges from 1.70 to 4.14 times, with average frequency ranked as cluster1, cluster3, cluster0, cluster4, and cluster2. The monetary value across the five clusters ranges

$K = 5$ clusters. The sum of the intracluster distances of five clusters was 4146.37 and 2290.15 when using the $K$-means and GCMPSO algorithms, respectively. The results revealed that the GCMPSO algorithm outperformed the $K$-means algorithm.
from NT$2,644.20 to NT$7,070.27 with the average monetary value for each donation given by each donor ranging from NT$1,668.97 to NT$1,945.26. The cluster ranking for monetary value is cluster1, cluster4, cluster0, cluster3, and cluster2. The results considering the RFM values imply that the donors in cluster1 make the highest contributions, and those in cluster2 make the least.

### E. CLUSTER GROUPING AND SEGMENT LABELS

#### 1) CLUSTER GROUPING

Similar clusters were grouped based on the RFM values of each cluster. In analyzing and comprising the segments, the frequency value should be considered first, followed by recency, and then monetary value. Based on this rule, we take the frequency of 3 within three years as the standard. One donation in one year is treated as the judgment of whether the donation periodicity is “significant”. Cluster1 (Frequency = 4.14) and cluster3 (Frequency = 3.43) have a frequency value above the average of 3, indicating significant donation periodicity. The donation frequency of cluster0 (Frequency = 2.23), cluster4 (Frequency = 2.18), and cluster2 (Frequency = 1.70) is less than 3 during the three years, which is not significant.

In the “significant” cluster, the recency of cluster1 and cluster3 is less than 365 (i.e., one year). The 266.52 days of cluster1 is less than the 292.29 days of cluster3. This indicates that donors in cluster1 and cluster3 are recent donors. Finally, the monetary value of Cluster1 is NT$7,070.27, which is higher than the NT$5,232.88 of cluster3. Cluster1 makes more contributions than cluster3. Therefore, cluster1 is classified as loyal donors, and cluster3 as active donors.

In the “less significant” cluster, the recency of clusters 0, 2, and 4 is more than 365 days. It was found that cluster0 made donations for three consecutive years (2016-2018); however cluster2 and cluster4 donated during 2016-2017, but made no donations in 2018. In addition, cluster0 (NT$3,810.91) has higher monetary value than cluster2 and cluster4, which have the last two rankings of donation amount among all clusters (NT$2,644.20 and NT$3,685.95). Therefore, cluster0 was classified as potential donors, and cluster2 and cluster4 were classified together as passive donors.

#### 2) DONOR SEGMENT LABELS

To present the characteristics of each segment, sociodemographic variables and RFM were considered simultaneously for labeling the groups. The first part of the name revealed the sociodemographic features, and the last part of the name represented the donation behavior. Loyal donors exhibited the highest RFM, followed by active donors, potential donors, and passive donors. Therefore, cluster1 was labeled as “male loyal donors,” cluster3 as “female active donors,” cluster0 as “potential donors,” and cluster2 and cluster4 were labeled together as “passive donors” who have not donated for a long time. The profiles of each donor segment are shown in Figure 6.

- **a: MALE LOYAL DONORS**

Cluster1 is labeled as male loyal donors. The segment contains 17% (N = 1,243) of the sample, and is more likely to be male. Most of the donors are between 30 and 40 years old, with a mean age of 40.83. The majority have a bachelor’s degree. Their donations were made in 2017 and 2018 but not in 2016. Most of the recency is over 300 days, with an average of 266.52 days, which is the least among the four segments. The frequency is mostly more than four times, with an average of 4.14 times, which was the highest among the four groups. In addition, the monetary value is NT$7,070.27, with a single donation of NT$1,945.26, which is the highest among the four groups. The monetary value of the total contributions by the cluster is NT$8,788,351. The male loyal donors are the smallest portion of all donors; however, they are the most important segment in comparison with the others. For fundraisers across all types of charity organizations, male loyal donors appear to be the single most feasible group on which to focus their efforts.

- **b: FEMALE ACTIVE DONORS**

Cluster3 is labeled female active donors. The second largest segment (N = 1,818, 24%), this is largely composed of women, typically 39.42 years of age, and with high education of a bachelor’s degree. Their donations occurred in 2017 and 2018 but not in 2016. Most of the recency ranges between 301 and 400 days, with an average of 292.29 days, which is second among the four segments. Most of the frequency ranges between 1.10 and 2.00 times, with an average of 3.43 times, which ranked second among the four groups. In addition, the monetary value is NT$5,232.88, with a single donation of NT$1,692.67. This group reports the highest amount of charitable donations (NT$9,513,370) and can be seen as a group where there is genuine loyal donor potential.

- **c: POTENTIAL DONORS**

Cluster0 is labeled potential donors, and comprises 1,542 individuals (21%). Most are women, aged 39.14 with a master’s degree; there are few men in this group. The educational achievement of a master’s degree is the highest across all segments. Donations occurred from 2016 to 2018.
Recency is mostly over 400 days, with an average of 570.15 days, which is ranked third among the four segments. Frequency mostly ranged between 1.10 and 2.00 times, with an average of 2.23 times, which ranked third among the four groups. In addition, the monetary value is NT$3,810.91, and the single donation of each donor in the segment is NT$1,783.87, higher than the female active donors (NT$1,692.67).
d: PASSIVE DONORS
Cluster4 and cluster2 were combined into the segment labeled passive donors. It is the largest of the four groups (N = 2,829), accounting for 38% of all samples. As the name suggests, this segment is characterized by low propensity to donate across all charity types. As such, they represent the least valuable group for fundraisers to target in general. The group is typically comprised of two types of individuals. First, mainly men, 40.71 years old, college degree, average donation recency ranked fourth (781.74 days), donation frequency ranked fourth (2.18 times), monetary value ranked fourth (NT$3,685.95). Second, mainly women, 36.5 years old, college degree, recency ranked fifth (823.14 days), donation frequency ranked fifth (1.70 times), monetary value ranked fifth (NT$2,644.20). The charitable donations occurred in 2016 and 2017, but not in 2018. The donors in this segment made no donations in more than a year until the end of 2018, implying that the donors in this segment have been inactive for a long time. The monetary value is the smallest among the four segments at NT$8,381,557 (3,188,350 and 5,193,207), thus comprising the less valuable donors in all four segments.

F. MARKETING STRATEGIES
Rather than having to formulate marketing strategies for all donors, the study formulates specific strategies for each segment based on the features of each. With the help of the experts in charitable donations, the suggested marketing strategies for each segment of donors are illustrated as follows.

1) MALE LOYAL DONORS
This is the most valuable among all segments according to the total monetary value. Therefore, the goal of the marketing strategies for this segment is to sustain donation motivation, maintain long-term relationships, and increase continuous donation behavior. Specifically, the donors should be encouraged to upgrade from one-time donation to continuous donors. The donors in this segment are mainly male, so the donation appeals could be designed to target the characteristics of males. Emotional appeal is one kind of advertising appeal, which stimulates the consumer’s emotional attitude towards the product by adding emotion to the advertisement [42]. Previous studies showed that males are triggered to donate by eliciting the emotion of pride [43]. Therefore, a “thank you certificate” should be issued for individuals after donation to initiate the pride of these male donors. In addition, men generally like charitable advertisements with self-help appeals [44]. The argument for the benefits of helping oneself could be presented in the charitable ads.

2) FEMALE ACTIVE DONORS
This segment is ranked second in terms of monetary value for each donor, but ranks first for the total amount of donations. Therefore, the marketing goal in this segment is to encourage more monetary value close to the male loyal donors. The frequency of the donation and the monetary value for single donations are expected to be upgraded. The donors in this segment are mainly female, and the donation message should therefore be designed to target female donors. A previous study showed that persuasive appeals that induced sympathy were more effective for women than men in encouraging donation behavior[43]. Sympathy promotes attention to the needs of others and caring. The negative images of beneficiaries are a useful trigger for women’s sympathy to make a donation [33]. In addition, the benefits of helping others could be emphasized in the fundraising projects [44] in order to promote more donations.

3) THE POTENTIAL DONOR
Upgrading to be active donors is the main goal of marketing strategies for the segment who are potential donors. The donation frequency is expected to be enhanced and to reach over two times per year. The opportunity for online donation platforms to contact donors should be increased in order to motivate the donors to engage in more charity-giving. Previous studies showed that the argument of charitable donation focusing on the statistical description of victims is more effective to increase the intention of charitable donations [45]. Most of the donors in the segment have high educational achievement of a Master’s degree, and are more knowledgeable and rational. Therefore, the charitable message with statistical descriptions could be delivered often to donors to initiate rational thinking about charity giving in order to increase the frequency of donations.

4) PASSIVE DONORS
This is the least profitable of all four segments. However, this segment should not be abandoned; in fact, more marketing resources should be allocated to these donors. The marketing goal of the segment is to awaken the donors in this segment to make charitable donations again. Such donors may be frequently informed of fundraising projects and motivated by more touching events. The donors in this segment are aged between 30 and 40, and so should have excellent technology skills such as internet and social media skills. Donors in this segment should be contacted by email, or invited to join the social media platform belonging to one non-profit organization, which should in turn trigger more opportunities for donation.

IV. DISCUSSION
This study has proposed a method combining RFM values and socio-demographic variables to segment donors using a PSO algorithm for clustering. Marketing strategies are also provided to enhance fundraising and maximize the monetary value. A case study was implemented by applying the method to analyze values of online charitable donors. Longitudinal online donation data were used instead of questionnaire survey data to analyze the profiles of online charitable donors in Taiwan. Longitudinal data compensate for the disadvantage
of “social desirability” when using self-reported questionnaires. The results show the whole profiles of online donors in Taiwan. The insights and implications of the study are illustrated as follows.

A. THE PROFILES OF GENERAL DONORS

From the demographic descriptions of all donors, some interesting phenomena were found. First, females were the main charitable donors on the online donation platform, accounting for 70%, consistent with previous studies [13], [14], [29]. Second, the majority of donors are aged between 30 and 40, which is the productive age and middle-income group in society, consistent with the finding of Kasri [15] who demonstrated the same phenomenon in the Islamic world. Donors aged over 60 are fewer, a trend that is consistent with previous findings suggesting that charitable donation tends to fall off after the age of retirement [15], [46]. Finally, 49% of all donors had a bachelor’s degree, indicating that the majority of donors are highly educated. This finding is consistent with the results of previous studies showing a strong positive correlation between education level and amounts of regular donation [18], [28], [30], [31].

B. THE INSIGHTS FOR EACH SEGMENT

Our results also provide some insights regarding each segment. Five clusters were found and four donor segments were categorized based on the similarity of the clusters. The segments of donors were labeled “male loyal donors,” “female active donors,” “potential donors” and “passive donors.” The male loyal group is the most profitable segment because the average frequency and monetary values are higher than the overall average value, which indicates that these donors give higher amounts of money frequently. We do not have to invest too many resources in sustaining a long-term relationship with this type of donor because the stable relationship has already been sustained for a considerable period of time. It is notable that the donors in the male loyal segment (cluster1) are mainly men who have the highest single donation across all groups (NT$1,945.26). This finding is consistent with previous studies, and indicates that the one-time monetary value of their donations is higher than that of women, although the number of men donating is less than that of women [29].

The second most profitable donors are the female active donors (cluster3), who have high frequency and monetary values above the average means. The donors in the female active segment are mainly women. This finding is consistent with previous studies, which demonstrated that females are the main contributors for charity fundraising [29]. The result implies that women have more empathy than men. Although the monetary value in this segment is less than that of the loyal donors, the total monetary value (NT$9,513,370) is the top of all segments. Managerial resources are expected to be allocated to this type of segment because of the fruitful amount of their donations. The potential donors (cluster0) with below average frequency and monetary values could have profit potential in the near future. Based on the analysis, this group has higher potential to become active donors in the long term. The platform should invest the majority of its resources in activating their charity donations and increasing the monetary value of their donations. Most of the donors in this segment have high educational achievement of a Master’s degree. Managers may consider posting more charity information with statistical descriptions to increase the opportunity of charity giving.

Passive donors (cluster2 and cluster4) constitute the largest of the four segments, and have the lowest recency, frequency, and monetary values compared to the other segments. In terms of recency, these donors had not donated on the online donation platform for more than one year, and in terms of frequency and monetary values, they do not donate a lot of money, even those who often donate. However, if we look at the monetary donation, cluster4 in the passive segment is not the lowest among all clusters (NT$3,685.95), implying the potential for high donation amounts. This cluster should be contacted again to resume their donations.

Finally, considering donation behavior by gender, the results showed that women who donate the largest donations are the main contributors to the online charitable platform. However, men’s single donation (NT$1,945.26 in cluster1; NT$1,904.78 in cluster4) is higher than that of women (NT$1,783.87 in cluster0; NT$1,668.97 in cluster2; NT$1,692.67 in cluster3). These results are consistent with previous studies, indicating that women donate more often, but make lower charitable donations than men [32], [33].

C. THE TECHNIQUE OF CLUSTERING

Clustering analysis can be used to reveal the structure inherent in the data. The well-known K-means algorithm is widely applied to practical issues. However, the K-means algorithm has some defects because the choice of its initialization pipeline and objective function of non-convex could fall into local optima. PSO has demonstrated outstanding performance in clustering in multi-dimensional spaces. However, the convergence speed is not fast enough when searching for the global optimal solution. Therefore, a GCM is used to improve PSO for detecting the better clusters amongst all data. The GCM is a quadratic transform and provides the continuous fractional expansion of numbers. It is similar to the shift transform corresponding to the quadratic iterative operator so that it allows a complete analysis of the qualitative and quantitative properties of chaos. The shift transformation can satisfy the dense periodic points, mixing and sensitivity of chaos. The characteristics and self-adjustment of Gaussian chaotic mapping can avoid PSO falling into local optimum, and enable it to deal with a particular data type.

D. MARKETING IMPLICATIONS

The findings discussed above have several practical implications. First, the donation platform should focus on relationship marketing that attempts to maintain relationships...
with donors. Relationship marketing is the key successful factor for fundraising upgrade, regardless of donor type [47], and is an effective approach to assist nonprofits in reducing the lapse rate of donors to their organization [48]. The concept of relationship marketing emphasizes the importance of developing long-term relationships with existing donors, and ensures that energy and resources are better spent on this group [4]. Numerous studies concerning relational marketing have argued that organizations should make more efforts with potential customers and ignore the less valuable customers [49]. Donors are different from consumers; they are motivated to spend money to help others, while consumers want to entertain themselves [6]. Logically, the marketing strategy applied to donors should be different from that applied to consumers. Therefore, we suggest that the online platform should allocate more resources to passive donors to awaken them again, especially for the cluster with high monetary donations (i.e., cluster4). These passive donors have a high probability of donation growth if there is an appropriate marketing strategy implemented with a focus on them.

The combination of socio-demographic variables and RFM values can help us better understand the profiles of donors. The online platform can put forward corresponding marketing strategies according to the characteristics of donors. Sex, education, and age are found to be salient features of each segment. Most of the donors on this donation platform are highly educated and middle-aged, between 30 and 40 years old, which is the productive working age. The channel of information delivery can be through emerging media, such as social networks [50]. In addition, the majority of donors are women; however, men’s single donation amount is higher than that of women. Contrary to intuition, men are donors who deserve to have their relationship maintained to raise their donation amount.

The report by netiCRM [2] showed that the proportion of online donations of small organizations with annual fundraising of less than NT$2 million has exceeded 30%, and the proportion of online donations of non-profit organizations was as high as 80% during 2018 in Taiwan. This phenomenon is of great significance. It reflects that most non-profit organizations use the Internet as their main communication channel, as it has become the new means of communication of ideas and recruitment of financial resources. The trend of making charitable donations via the Internet has emerged, and exploring the behavior mode of the online donors has been an important issue in Taiwan. Although the study provides several implications and contributions for the literature and charity marketing managers, there are some limitations that should be noted. First, the dataset came with some limitations; specifically, it was limited to the analysis of 2016-2018 donation records, so the donor segmentation is not suitable for inferring other years. Besides, some columns of the donation records with missing data were removed, so the dataset for analysis was not the original one. Since donors contributed in 2016-2017, their age was calculated based on the average age of three years.

It should be noted that this is a general study aimed at providing a picture of the online charitable donors in Taiwan. Thus, it must be interpreted within the appropriate context and considering its limitations. Further studies are needed to investigate more issues, such as the giving behaviors of different types of donors, and the giving behaviors in western countries or case studies of specific nonprofit organizations. Similar studies could be considered to employ more donation data and psychological factors (i.e., empathy, moral norms, and social responsibility) of charity giving to gain a thorough understanding of online donors. Such studies would contribute towards a further understanding of marketing and management practices of online donation platforms in Taiwan.

REFERENCES

[1] A. Dogan, E. Calik, and F. Calisir, “Organizational factors affecting individuals to donate to NPOs in the Turkish context," VOLUNTAS, Int. J. Voluntary Nonprofit Organizations, vol. 2020, pp. 1–13, Feb. 2020.
[2] netiCRM. (2020). What is the Overall Appearance of the Variable 2018 Donation in Taiwan. [Online]. Available: https://neticrm.tw/report/2018
[3] N. J. D. Vries, R. Reis, and P. Moscati, “Clustering consumers based on trust, confidence and giving behaviour: Data-driven model building for charitable involvement in the Australian not-for-profit sector,” PLoS ONE, vol. 10, no. 4, Apr. 2015, Art. no. e0122133.
[4] S. Boenigk and C. Scherhag, “Effects of donor priority strategy on relationship fundraising outcomes,” Nonprofit Management, Leadership, vol. 24, no. 3, pp. 307–336, Mar. 2014.
[5] C. H. Lovelock, and C. B. Weinberg, Marketing for Public and Nonprofit Managers. New York, NY, USA: Wiley, 1984.
[6] C. Rupp, S. Kern, and B. Helmig, “Segmenting nonprofit stakeholders to enable successful relationship marketing: A review,” Int. J. Voluntary Nonprofit Sector Marketing, vol. 19, no. 2, pp. 76–91, May 2014.
[7] H. Bo Liu, B. McCarthy, T. Chen, S. Guo, and X. Song, “The Chinese wine market: A market segmentation study,” Asia Pacific J. Marketing Logistics, vol. 26, no. 3, pp. 450–471, Jun. 2014.
[8] P. Sultan, H. Y. Wong, and M. Sigala, “Segmenting the Australian organic food consumer market,” Asia Pacific J. Marketing Logistics, vol. 30, no. 1, pp. 153–181, Jan. 2018.
[9] C.-H. Cheng and Y.-S. Chen, “Classification of the segmentation of customer value via RFM model and RS theory,” Expert Syst. Appl., vol. 36, no. 3, pp. 4176–4184, Apr. 2009.
[10] H. Khobzi, E. Akhondzadeh-Noughabi, and B. Minaei-Bidgoli, “A new application of RFM clustering for guild segmentation to mine the pattern of using banks’ e-Payment services,” J. Global Marketing, vol. 27, no. 3, pp. 178–190, May 2014.
[11] J. Orange, S. Sharif, and C. Uviedo, “Using RFM segmentation to support alumni reunion fundraising,” J. Educ. Advancement Marketing, vol. 3, no. 4, pp. 325–339, 2019.
[12] A. Hamzehei, M. Fathian, H. Farvaresh, and M. R. Gholamian, “A new methodology to study customer electrocardiogram using RFM analysis and clustering,” Manage. Sci. Lett., vol. 1, no. 2, pp. 139–148, Apr. 2011.
[13] L. Shelley and M. J. Polonsky, “Do charitable causes need to segment their current donor base on demographic factors?: An Australian examination,” Int. J. Nonprofit Voluntary Sector Marketing, vol. 7, no. 1, pp. 19–29, Feb. 2002.
[14] Y.-K. Lee and C.-T. Chang, “Who gives what to charity? Characteristics affecting donation behavior,” Social Behav. Personality, Int. J., vol. 35, no. 9, pp. 1173–1180, Jan. 2007.
[15] R. A. Kasi, “Giving behaviors in Indonesia: Motives and marketing implications for Islamic charities,” J. Islamic Marketing, vol. 4, no. 3, pp. 306–324, Sep. 2013.
[16] K. J. Smka, R. Grohs, and I. Eckler, “Increasing fundraising efficiency by segmenting donors,” Australas. Marketing J., vol. 11, no. 1, pp. 70–86, May 2003.
P. Anitha and M. M. Patil, “RFM model for customer purchase behavior,” J. Nonprofit Public Sector Marketing, vol. 22, no. 1, pp. 38–54, Feb. 2010.

L. Y. Chuang, Y. D. Lin, and C. H. Yang, “Data clustering using chaotic particle swarm optimization,” IAENG Int. J. Comput. Sci., vol. 39, no. 2, pp. 208–213, 2012.

L. Y. Chuang, Y. D. Lin, and C. H. Yang, “An improved particle swarm optimization for data clustering,” in Proc. Int. MultiConf. Eng. Comput. Scientists (IMECS), Hong Kong, 2012, pp. 440–445.

S. Z. Selim and M. A. Ismail, “K-means-type algorithms: A generalized convergence theorem and characterization of local optimality,” IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-6, no. 1, pp. 81–87, Jan. 1984.

M. Mittal, L. M. Goyal, D. J. Hemanth, and J. K. Sethi, “Cluster-approaching techniques for high-dimensional databases: A review,” Wiley Interdiscip. Rev., Data Mining Knowl. Discovery, vol. 9, no. 3, p. e1300, May 2019.

C.-H. Yang, L.-Y. Chuang, and Y.-D. Lin, “Epistasis analysis using an improved fuzzy C-means-based entropy approach,” IEEE Trans. Fuzzy Syst., vol. 28, no. 4, pp. 718–730, Apr. 2020.

P. Anitha and M. M. Patil, “RFM model for customer purchase behavior using K-means algorithm,” J. King Saud Univ.-Comput. Inf. Sci., vol. 28, no. 4, pp. 718–730, Apr. 2020.

S. Monalisa and F. Kurnia, “Analysis of DBSCAN and K-means algorithm for evaluating outlier on RFM model of customer behaviour,” Int. J. Advertising, vol. 42, no. 1, pp. 69–79, Jan. 2013.

F. F. Brunel and M. R. Nelson, “Explaining gendered responses to ‘help-’self’ and ‘help-others’ charity ad appeals: The mediating role of worldviews,” J. Advertising, vol. 29, no. 3, pp. 15–28, 2000.

N. Kim, “Advertising strategies for charities: Promoting consumers’ donation of time versus money,” Int. J. Advertising, vol. 33, no. 4, pp. 707–724, Jul. 2014.

A. Hughes, Strategic Database Marketing, Chicago, IL, USA: Probus, 1994.

J. Y.-C. Hsu, “Research on customer segmentation model by clustering,” in Proc. 7th Int. Conf. Electron. Commerce (ICEC), New York, NY, USA, 2005, pp. 316–318.
TZER-SHYONG CHEN (Member, IEEE) received the Ph.D. degree in computer science from the Department of Electrical Engineering, National Taiwan University, Taiwan, in 1996. He is currently the Department Chair of information management with Tunghai University, Taiwan. He served on an evaluation committee of the Institute of Electrical Engineering Taiwan. He has authored/co-authored over 80 refereed publications. His main research interests include information security, cryptography, and network security.

TE-YI CHANG is currently a Professor with the Graduate School of Graduate Institute of Tourism Management, National Kaohsiung University of Hospitality and Tourism. Her research interests include tourism behavior, human resources management and performance management, with those findings published in journals, such as Tourism Management (SSCI), Journal of Travel Research (SSCI), JoHLSTE (SSCI), Revista Internacional de Sociologia (SSCI), and Cornell Hospitality Quarterly (SSCI).

YU-DA LIN (Senior Member, IEEE) received the M.S. and Ph.D. degrees in computer science from the Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan, in 2011 and 2015, respectively. He is currently an Assistant Professor with the Department of Computer Science and Information Engineering, National Penghu University of Science and Technology. He has authored/co-authored more than 100 refereed publications. His main research interests include artificial intelligence, biomedical informatics, bioinformatics, and computational biology.

Dr. Lin is a Senior Member of the IEEE Tainan Section, IEEE Young Professionals, and IEEE Computational Intelligence Society Membership. He is also an Associate Editor of IEEE Access Journal.

---

C.-W. Hsu et al.: Who Donates on Line? Segmentation Analysis and Marketing Strategies Based on Machine Learning