Available online at www.sciencedirect.com
ScienceDirect
IFAC PapersOnLine 51-11 (2018) 1174–1179

Clustering for inventory control systems

E. Balugani*. F. Lolli.**
R. Gamberini***. B. Rimini****. A. Regattieri*****

* *Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Modena, Italy (e-mail: elia.balugani@unimore.it).*
** *Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Modena, Italy (e-mail: francesco.lolli@unimore.it).*
*** *Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Modena, Italy (e-mail: rita.gamberini@unimore.it).*
**** *Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Modena, Italy (e-mail: bianca.rimini@unimore.it).*
******University of Bologna, Bologna, Italy (e-mail: alberto.regattieri@unibo.it).*

Abstract: Inventory control is one of the main activities in industrial plant management. Both process owners and line workers interact daily with stocks of components and finite products, and an effective management of these inventory levels is a key factor in an efficient manufacturing process. In this paper the algorithms k-means and Ward's method are used to cluster items into homogenous groups to be managed with uniform inventory control policies. This unsupervised step reduces the need for computationally expensive inventory system control simulations. The performance of this methodology was found to be significant but was strongly impacted by the intermediate feature transformation processes.

Keywords: inventory control, clustering, k-means, ward's method, intermittent demand, spare parts, machine learning, simulation.

1. INTRODUCTION

Inventory control is the act of supervising stocked components and finite products. In a production environment, both plant managers and operators interact daily with stocked items. They rely on automated inventory control systems that manage the amount of items to order or produce given the historical demand, the production schedule and the current stock levels.

This paper focuses on components management. The proposed methodology is designed to deal with inventory control systems, managing both intermittent and non-intermittent items. Intermittent items, e.g. spare parts (Regattieri *et al.*, 2005), are characterized by frequent periods with no demand and high variability in the demand size when a demand takes place (Croston, 1972; Syntetos, Boylan and Croston, 2005).

Forecasting for intermittent items usually requires ad-hoc estimators. The most commonly used methodology in this regard is the Syntetos-Boylan Approximation (SBA) (Syntetos and Boylan, 2005), a modification of the standard exponential smoothing procedure. Neural network methods (Lolli, Gamberini, *et al.*, 2017) and SARIMA techniques (Gamberini *et al.*, 2010; Lolli *et al.*, 2011; Lolli, Gamberini, Regattieri and Rimini, 2014; Lolli, Gamberini, Regattieri, Rimini, *et al.*, 2014) alternative to the SBA have proven to be highly effective in dealing with items characterized by intermittent consumption.

The inventory control systems for intermittent demand differ from those for traditional items. While traditional methodologies can be adapted (Gamberini *et al.*, 2014), ad hoc inventory control systems (Teunter, Syntetos and Babai, 2010; Babai, Jemai and Dallery, 2011) can better leverage on the characteristics of intermittent demand. Accounting for these characteristics reduces the generality of the proposed methodologies, but it enables environments to be modelled that are subject to both intermittent consumption and stock wear out issues such as obsolescence and perishability (Balugani *et al.*, 2017).

When the inventory control system under analysis is difficult to model, as in the intermittent demand case, simulation techniques can be exploited to optimize the implemented policies. The historical series of the items are simulated multiple times with different settings to identify the best inventory control system parameters. This simulation phase can be very computationally expensive, thus supervised classification techniques are implemented to reduce the simulation workload to a subset of items. Both decision trees (Lolli, Ishizaka, *et al.*, 2017) and AHP approaches (Lolli, Ishizaka and Gamberini, 2014) have been successfully used for this purpose.

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2018.08.431

The aim of this paper is to further reduce the computational workload by implementing clustering techniques and dividing the items in a non-supervised fashion before the simulation occurs. This initial subdivision allows for the simulation of just a limited number of items in each cluster, with considerable time savings.

Section 2 defines the inventory control system used in the paper. Section 3 introduces the clustering algorithms implemented. Section 4 outlines the proposed methodology. Section 5 defines the experiment designed to validate the proposed methodology and outlines the obtained results. Section 6 presents the conclusions and the future research agenda.

2. INVENTORY CONTROL SYSTEM

Periodic inventory control system

The inventory control system used in this paper is periodic: an order for new items is placed each T period and requires L periods to arrive. This policy evaluates both unitary holding costs c_h and unitary order costs c_o, the holding costs are paid each period per stocked unit, while the order costs are paid whenever an order is placed. The average total cost per period c_t is:

$$c_t = c_h \frac{Q}{2} + \frac{c_o}{T} \tag{1}$$

Where Q is the order up to a level, and can be written as:

$$Q = D \cdot p \cdot (T + L) \tag{2}$$

Where D is the expected positive demand (with standard deviation σ) for a single period, and p is the probability a positive demand occurs. This notation can model both intermittent and non-intermittent items.

(1) can be derived and equated to zero to minimize c_t, thereby obtaining T_{best} as the optimal value for T:

$$T_{best} = \sqrt{\frac{c_o}{c_h} \cdot \frac{2}{D \cdot p}} \tag{3}$$

As outlined in (1) the model does not consider safety stock costs, if a safety stock is required, it is defined after the optimization of c_t. The safety stock requirements can be expressed in terms of target csl, the probability of there being a stockout during $T + L$ periods. A value of csl equal to 0.5 leads to no safety stock.

Simulation

While the inventory control system analysed in Section 2.1 can be optimized analytically, the proposed methodology outlined in Section 4, is general and can be also applied to policies that are analytically unsolvable. From a different perspective, even with analytically tractable inventory control systems, real items might not behave exactly like the hypothesized theoretical model since their T_{best} could be shifted.

Both these issues can be solved by exhaustive simulation. The historical series of an item is simulated multiple times with the chosen periodic inventory control system and different values of T. The simulation leading to the lowest empirical value of c_t defines T_{best}. The only drawback of simulation is that it is time consuming , particularly the simulation of multiple values of T for each item. The methodology presented in Section 4 deals with this limitation.

3. CLUSTERING ALGORITHMS

3.1 K-means

The K-means algorithm (MacQueen, 1967) is one of the most commonly-used methods for unsupervised classification.

Given a set E_N of N item features and a predetermined number of clusters K the algorithm:

1. Randomly initializes K cluster centres $\hat{\boldsymbol{x}}_{\boldsymbol{k}}$.
2. Assigns each item, with features $\boldsymbol{x_i}$, to the nearest cluster C_k using Euclidian distances:

$$C_k = \begin{Bmatrix} \boldsymbol{x_i}: \|\boldsymbol{x_i} - \hat{\boldsymbol{x}}_{\boldsymbol{k}}\| \leq \|\boldsymbol{x_i} - \hat{\boldsymbol{x}}_{\boldsymbol{j}}\|, \\ \boldsymbol{x_i} \in E_N, \forall j \epsilon \{1, \dots, K\} \end{Bmatrix} \tag{4}$$

3. If the iteration is not the first one and at least one item changed cluster in step 2, the cluster centres are recalculated:

$$\hat{\boldsymbol{x}}_{\boldsymbol{k}} = \sum_{x_i \in E_N} \boldsymbol{x_i} \tag{5}$$

When cluster centre has been located in its items' mean, the algorithm then moves to step 2.

If no item changed cluster in step 2, the algorithm terminates.

The K-means is guaranteed to converge. However, different choices for the initial cluster centres can lead to different end results. In order to weight all the features fairly, the K-means is used on normalized features. The algorithm produces spherical clusters and linear divisions between clusters in the features space.

3.2 Ward's method

Ward's method (Ward, 1963) is a hierarchical clustering algorithm that minimizes the sum of square errors within clusters.

Given a set E_N of N item features, the method:

1. Initializes a cluster for each item and defines $\hat{\boldsymbol{x}}_{\boldsymbol{k}}$ initial cluster centres as:

$$\hat{\boldsymbol{x}}_{\boldsymbol{k}} = \boldsymbol{x_k} \quad \forall \boldsymbol{x_k} \in E_N \tag{6}$$

2. Calculates the sum of squared errors between clusters as:

$$Err(\hat{\boldsymbol{x}}_{\boldsymbol{i}}, \hat{\boldsymbol{x}}_{\boldsymbol{j}}) = \|\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \hat{\boldsymbol{x}}_{\boldsymbol{j}}\|^2 \quad \forall i, j \in \{1, \dots, K\} \tag{7}$$

Merges the clusters with the minimum sum of squared errors between and calculates the new cluster centre:

$$\hat{\boldsymbol{x}}_{\boldsymbol{k}} = \sum_{x_i \in C_k} \boldsymbol{x_i} \tag{8}$$

This step can be simplified using a recursive algorithm (Lance and Williams, 1967), in this case an explicit calculation of the new $\hat{x}_k$ is not required.

Step 2 is repeated until the desired number of clusters K is reached.

The minimization of the sum of squared errors between merged clusters coincides with the minimization of the sum of squared errors within the remaining clusters, and the sum of squared errors of the system is fixed. As in the K-means case, the algorithm produces spherical clusters and linear divisions between clusters in the features space.

4. PROPOSED METHODOLOGY

The clustering algorithms presented in Section 3 are applied to a set of items following a defined inventory control system. The inventory control system can be the one outlined in Section 2.1 but, as explained in Section 2.2, this methodology is general.

The items are not exhaustively simulated as in Section 2.2, instead the clustering algorithms operate on the items' features by dividing them into clusters with homogeneous T_{best}. The underlying hypothesis is that clusters obtained using meaningful features in terms of an inventory control system can be used to define inventory control system policies. In practice, if the items fall into clusters with mostly homogeneous T_{best}, it is possible to characterize each cluster afterwards by simulating just a limited number of items, which thereby decreases the overall simulation time. Once a T_{best} is assigned to a cluster, the items within the cluster are managed with an optimized periodic inventory control system with a T_{best} review time.

When an analytical solution is available but the data do not follow the theory exactly, a theory-based feature transformation might be advantageous to simplify the feature space for the clustering algorithms. The policy outlined in Section 2.1, for example, can be rewritten in spherical coordinates:

$$r = \sqrt{c_o^2 + (c_h \cdot D \cdot p)^2} \tag{9}$$

$$\alpha = arctan\left(\frac{c_o}{c_h \cdot D \cdot p}\right) \tag{10}$$

$$T_{best} = \sqrt{2} \cdot \sqrt{tan(\alpha)} \tag{11}$$

In this coordinate system, T_{best} is a linear function of the single variable $\sqrt{tan(\alpha)}$.

5. EXPERIMENTAL SETTING AND RESULTS

In this experiment 625 items are generated by equally spacing 25 values of r, 25 values of $\sqrt{tan(\alpha)}$ and computing all their combinations. The extreme values for both features are listed in Table 1, the limits of $\sqrt{tan(\alpha)}$ are chosen to ensure theoretical values of T_{best} equally spaced between 2 and 8. A value of c_h and p is randomly assigned to each item from the intervals in Table 1.

Table 1. Simulated feature ranges

		min	max
Fixed features	L	2	
	csl	0.5	
	$\frac{\sigma}{D}$	0.5	
Equally spaced features	r	100	200
	$\sqrt{tan(\alpha)}$	1.35	5.70
Randomly generated features	c_h	0.1	0.6
	p	0.5	0.7

The remaining features needed for simulation and clustering are computed by inverting (6) and (7):

$$c_o = r \cdot cos(\alpha) \tag{12}$$

$$D = \frac{r}{D \cdot p} \cdot sin(\alpha) \tag{13}$$

Each item is assigned the same csl, L, and relative standard deviation $\frac{\sigma}{D}$, as defined in Table 1.

From the features p and D, an intermittent historical series of 200 periods is generated for each item. Each period has probability p of presenting a positive demand and, when a positive demand takes place, its size is normally distributed with expected value D and standard deviation σ. All the historical series are simulated three times, using the inventory control system outlined in Section 2 and values of T equal to 3, 5 and 7 respectively. After the simulations, the items are assigned to three homogenous classes according to the value of T minimizing their empirical c_t. The value of T of each class is an estimate its items' T_{best}.

After the division into classes, the overall number of items is reduced. The clustering algorithms in Section 3 group the items by similarity in the features space and divide them where they are dissimilar. The generated items are homogeneously spaced in the spherical coordinates, thus there is no clear division between clusters. From a different perspective, only three integer values of T are simulated, while most items could benefit from intermediate real values of T_{best}. As a consequence, the difference in the simulated c_t for two subsequent T is sizable for some items and small for others, and this difference is measured as:

$$\Delta c_t = \begin{cases} |c_t - c_{t2}| & if\ T_{best} \epsilon \{3,7\} \\ min\{|c_t - c_{t1}|, |c_t - c_{t3}|\} & if\ T_{best} = 5 \end{cases} \tag{14}$$

Where c_{t1} is the empirical average total cost per period when the item is simulated with $T = 1$, c_{t2} is the cost when $T = 2$, and c_{t1} is the cost when $T = 3$. Only 25% of items per class with the highest Δc_t are kept in order to differentiate the classes and retain only the most meaningful items.

The selected items are clustered with three different sets of features using both the K-means and Ward's method, the clustering algorithms are not provided with the simulated T_{best}. The three sets of features are designed as follows:

1. All the inventory control system features (c_o, c_h, D, $\frac{\sigma}{D}$, p, L, csl), normalized between 0 and 1.
2. The two combined features c_o and $c_h \cdot D \cdot p$, normalized between 0 and 1.
3. The transformed features r and $\sqrt{tan(\alpha)}$, normalized between 0 and 1.

The results for the three sets are reported respectively in Tables 2, 3, and 4.

Figures 1 and 2 plot the clustering obtained using the second set of features for the K-means and Ward's method respectively. Figure 3 plots the clustering obtained using the third set of features since the solution for both algorithms is the same.

Table 2. All features results

	T_{best}	3	5	7
K-means	Precision	0.91	0.41	0.54
	Recall	0.59	0.57	0.54
	F-measure	0.72	0.48	0.54
Ward's method	Precision	0.95	0.42	0.58
	Recall	0.70	0.45	0.69
	F-measure	0.81	0.43	0.63

Table 3. Combined features results

	T_{best}	3	5	7
K-means	Precision	1	0.46	0.52
	Recall	0.83	0.49	0.58
	F-measure	0.91	0.48	0.55
Ward's method	Precision	0.98	0.49	0.57
	Recall	0.94	0.73	0.33
	F-measure	0.96	0.59	0.41

Table 4. Transformed features results

	T_{best}	3	5	7
K-means	Precision	0.98	0.84	0.93
	Recall	1	0.92	0.83
	F-measure	0.99	0.88	0.88
Ward's method	Precision	0.98	0.84	0.93
	Recall	1	0.92	0.83
	F-measure	0.99	0.88	0.88

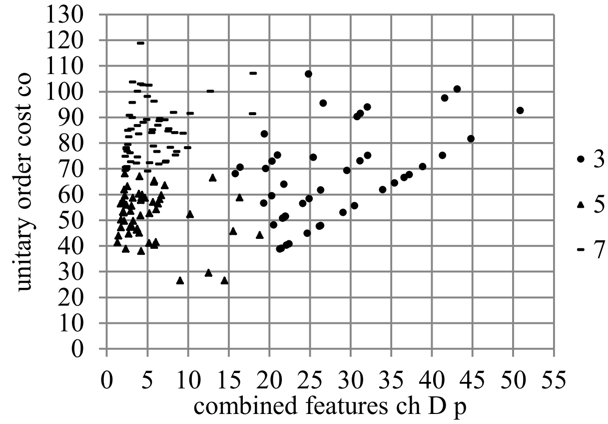

Fig. 1. K-means clustering obtained using of the second set of features.

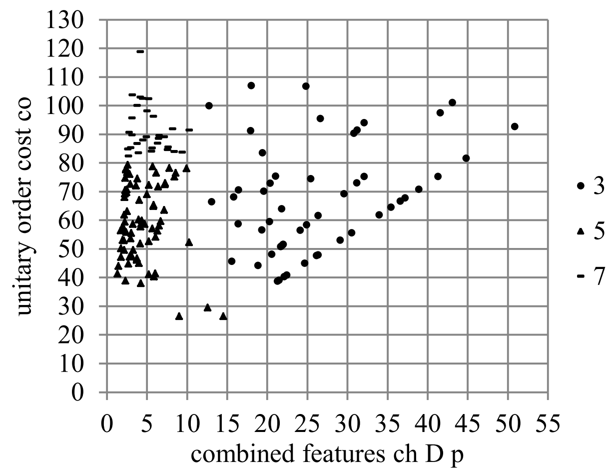

Fig. 2. Ward's method clustering obtained using the second set of features.

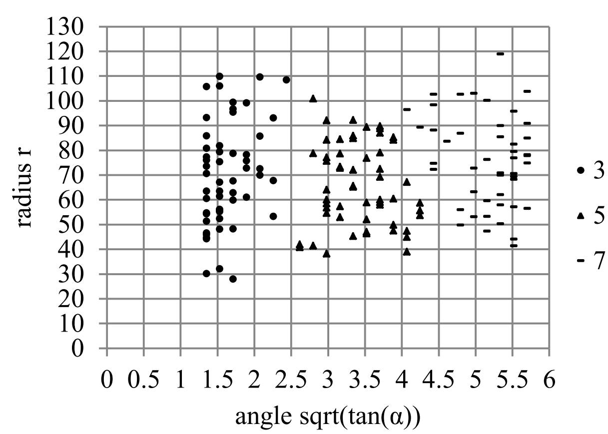

Fig. 3. Clustering obtained using of the third set of features.

In order to capture the quality of the solutions, three widely-used performance measures (Fawcett, 2006) are computed for each experiment, clustering algorithm, and cluster:

1. Precision, ratio between the number of items correctly classified in the cluster and the total number of items in the cluster.
2. Recall, ratio between the number of items correctly classified in the cluster and the total number of items that are assumed to belong to the cluster.
3. F-measure, harmonic mean of precision and recall.

Using the first set of features and the K-means algorithm the cluster $T_{best} = 3$ has high precision and small recall; the cluster is homogeneous but contains just a fraction of the items belonging to class $T_{best} = 3$. The remaining two clusters have low levels of both precision and recall. They contain a non-homogenous mix of items from all the three classes.

Ward's method performs better than the K-means using the first set of features, the recall for cluster $T_{best} = 3$ is significantly higher with 70% of the items of class $T_{best} = 3$ packed together. The cluster $T_{best} = 7$ increases its recall performance, contains 69% of the items of class $T_{best} = 7$, but is polluted by items from class $T_{best} = 5$ with a precision of 0.58. The cluster $T_{best} = 5$ is, as in the previous case, a non-homogenous mix of items from different classes.

Using the second set of features and the K-means algorithm, the cluster $T_{best} = 3$ has both high precision and recall. The cluster is homogeneous and contains 83% of the items of class $T_{best} = 3$. The remaining two clusters have a low level of both precision and recall. They contain a non-homogeneous mix of items from class $T_{best} = 5$ and class $T_{best} = 7$.

Ward's method performs better than the K-means using the second set of features. The recall for cluster $T_{best} = 3$ is higher, packing 94% of the items of class $T_{best} = 3$ the cluster near completely separates these items from the others. The cluster $T_{best} = 5$ increases its recall while remaining polluted by elements of class $T_{best} = 7$, while the recall of cluster $T_{best} = 7$ decreases sharply.

Overall, the use of second set of features seems slightly more effective than the first one. Nevertheless, as reported in Figs. 1 and 2, both clustering algorithms define spherical linearly divisible clusters in the features space, while a correct classification should refer to the slope of the items. As described in Section 3, this is the intrinsic behaviour of these clustering algorithms.

The third set of features overcomes the linearity problem with a transformation that generates linearly separable classes. As a result, the precision and recall for all the clusters are high, the items are well separated into homogenous groups, as shown in Fig. 3. In the last scenario the solutions for both the K-means and Ward's method are the same.

6. CONCLUSIONS

The use of the K-means and Ward's method is a viable substitute for inventory control system simulation when the clustering features are not directly obtained from the inventory control system. Theoretically grounded feature transformations are needed in order to obtain spherical linearly-divisible clusters in the features space.

Future research could apply non-linear and thus non-spherical clustering algorithms (e.g. spectral clustering) to overcome the need for theoretically grounded feature transformations, thus extending the methodology to inventory control systems that lack a strong analytical background.

REFERENCES

Babai, Z. M., Jemai, Z. and Dallery, Y. (2011) 'Analysis of order-up-to-level inventory systems with compound Poisson demand', *European Journal of Operational Research*, 210(3), pp. 552–558.

Balugani, E. *et al.* (2017) 'Inventory control system for intermittent items with perishability', in *24th International conference on Production Research*. Poznan.

Croston, J. D. (1972) 'Forecasting and Stock Control for Intermittent Demands', *Operational Research Quarterly (1970-1977)*, 23(3), pp. 289–303.

Fawcett, T. (2006) 'An introduction to ROC analysis', *Pattern Recognition Letters*, 27(8), pp. 861–874.

Gamberini, R. *et al.* (2010) 'Forecasting of Sporadic Demand Patterns with Seasonality and Trend Components: An Empirical Comparison between Holt-Winters and (S)ARIMA Methods', *Mathematical Problems in Engineering*, 2010, pp. 1–14.

Gamberini, R. *et al.* (2014) 'Dynamic Re-order Policies for Irregular and Sporadic Demand Profiles', *Procedia Engineering*, 69, pp. 1420–1429.

Lance, G. N. and Williams, W. T. (1967) 'A General Theory of Classificatory Sorting Strategies: 1. Hierarchical Systems', *The Computer Journal*, 9(4), pp. 373–380.

Lolli, F. *et al.* (2011) 'Application of tramo-seats automatic procedure for forecasting sporadic and irregular demand patterns with seasonality', in *13th Int. Conference on Harbor, Maritime and Multimodal Logistics Modeling and Simulation, HMS 2011, Held at the International Mediterranean and Latin American Modeling Multiconference, I3M 2011*, pp. 214–220.

Lolli, F., Gamberini, R., Regattieri, A. and Rimini, B. (2014) 'Application of Tramo-seats automatic procedure for forecasting sporadic and irregular demand patterns', in *OPT-i 2014 - 1st International Conference on Engineering and Applied Sciences Optimization, Proceedings*, pp. 1435–1445.

Lolli, F., Gamberini, R., Regattieri, A., Rimini, B., *et al.* (2014) 'On the Liaison between forecasting and periodic inventory control: Application of Sarima models to intermittent demand profiles', in *OPT-i 2014 - 1st International Conference on Engineering and Applied Sciences Optimization, Proceedings*, pp. 1426–1434.

Lolli, F., Ishizaka, A., *et al.* (2017) 'Decision Trees for Supervised Multi-criteria Inventory Classification', *Procedia Manufacturing*. Modena, 11, pp. 1871–1881.

Lolli, F., Gamberini, R., *et al.* (2017) 'Single-hidden layer neural networks for forecasting intermittent demand', *International Journal of Production Economics*. Elsevier, 183(May 2015), pp. 116–128.

Lolli, F., Ishizaka, A. and Gamberini, R. (2014) 'New AHP-based approaches for multi-criteria inventory classification', *International Journal of Production Economics*. Elsevier, 156, pp. 62–74.

MacQueen, J. (1967) 'Some methods for classification and analysis of multivariate observations', *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1, pp. 281–297.

Regattieri, A. *et al.* (2005) 'Managing lumpy demand for aircraft spare parts', *Journal of Air Transport Management*, 11(6), pp. 426–431.

Syntetos, A. A. and Boylan, J. E. (2005) 'The accuracy of intermittent demand estimates', *International Journal of Forecasting*, 21(2), pp. 303–314.

Syntetos, A. A., Boylan, J. E. and Croston, J. D. (2005) 'On the categorization of demand patterns', *Journal of the Operational Research Society*, 56(5), pp. 495–503.

Teunter, R. H., Syntetos, A. A. and Babai, Z. M. (2010) 'Determining order-up-to levels under periodic review for compound binomial (intermittent) demand', *European Journal of Operational Research*. Elsevier B.V., 203(3), pp. 619–624.

Ward, J. H. (1963) 'Hierarchical Grouping to Optimize an Objective Function', *Journal of the American Statistical Association*, 58(301), p. 236.