

CLASSIFICATION OF LIQUIDS USING A PATCH ANTENNA AND HIERARCHICAL CLUSTERING ALGORITHMS

Ebru EFEÖĞLU, Department of Software Engineering, Kütahya Dumlupınar University, Turkey,
ebru.efeoglu@dpu.edu.tr

( <https://orcid.org/0000-0001-5444-6647>)

Gürkan TUNA*, Department of Computer Programming, Trakya University, Turkey,
gurkantuna@trakya.edu.tr

( <https://orcid.org/0000-0002-6466-4696>)

Received: 11.05.2021, Accepted: 23.10.2021

*Corresponding author

Research Article

DOI: 10.22531/muglajsci.936406

Abstract

Detection of hazardous liquids used in explosive production is important in terms of public safety and health. Because many threats can be prevented by detecting these liquids at security controls points. As existing methods have some disadvantages in terms of accuracy, practicality or reliability, there is a demand for new methods for hazardous liquid detection. In this paper, a circular patch antenna for hazardous liquid detection was designed and by connecting to a vector network analyzer a group of measurements was made. Then, this dataset was used by hierarchical clustering algorithms employed in this study to detect hazardous liquids. The results show that high classification accuracy can be achieved when Ward linkage method is preferred.

Keywords: Distance metrics, Hazardous liquids, Hierarchical clustering, Linkage methods, Patch antenna

YAMA ANTEN VE HİYERARŞİK KÜMELEME ALGORİTMALARI KULLANILARAK SIVILARIN SINIFLANDIRILMASI

Özet

Patlayıcı üretiminde kullanılan tehlikeli sıvıların tespiti kamu güvenliği ve sağlığı açısından önemlidir. Çünkü güvenlik kontrol noktalarında bu sıvılar tespit edilerek birçok tehdidin önüne geçilebilmektedir. Mevcut yöntemlerin doğruluk, pratiklik veya güvenilirlik açısından bazı dezavantajları olduğundan, tehlikeli sıvı tespiti için yeni yöntemlere talep vardır. Bu makalede, tehlikeli sıvı tespiti için dairesel bir yama anteni tasarlanmış ve bir vektör ağ analizörüne bağlanarak bir grup ölçüm yapılmıştır. Daha sonra bu veri kümesi, tehlikeli sıvıları tespit etmek için bu çalışmada kullanılan hiyerarşik kümeleme algoritmaları tarafından kullanılmıştır. Sonuçlar, Ward bağlantı yöntemi tercih edildiğinde yüksek sınıflandırma doğruluğu elde edilebileceğini göstermektedir.

Anahtar Kelimeler: Mesafe metrikleri, Tehlikeli sıvılar, Hiyerarşik kümeleme, Bağlantı yöntemleri, Yama anten

Cite

Efeöğlü, E., Tuna, G., (2021). "Classification of Liquids Using a Patch Antenna and Hierarchical Clustering Algorithms", *Mugla Journal of Science and Technology*, 7(2), 40-44.

1. Introduction

In recent years, explosives detection has become an important problem for counter-terrorism reasons [1,2]. Therefore, studies are carried out to develop faster and more sensitive methods for explosive detection [3,4]. For this purpose, the use of Terahertz time domain spectroscopy [5], Raman spectroscopy [6] and Raman-based standoff detection systems [7], Nuclear magnetic resonance [8], infrared spectroscopy [9] and X ray systems [10] have been proposed. However, each of these techniques has some specific limitations. Therefore, the search for more practical and reliable methods for explosive detection still continues.

Each liquid has a different chemical composition and dielectric permeability value. When an electromagnetic wave sent to a liquid, it creates an interaction between the molecules of the liquid. As a result, the molecules rotate and align in the electromagnetic field. In this alignment process, the speed of the wave and the magnitude of the wave decrease. The signal amplitude reflected from the liquid is a function of the change in the dielectric constant and dielectric loss factor of the liquid [11]. Therefore, there will be variation in the reflection coefficients of liquids and hazardous liquids can be detected using this difference in the reflection coefficient of liquids. Microwave methods are commonly used to measure the dielectric properties of liquids. A resonator has been designed to characterize the dielectric properties of liquids [12,13]. It was used to

determine the alcohol concentration in alcoholic liquids [14]. Machine learning techniques, which have a wide range of usage and are applied for different classification goals, are used in the classification of hazardous liquids, too. For instance, liquid flammability ratings were estimated by considering aerosolization using a hierarchical clustering method [15]. K-means cluster analysis was used for the prevention of hazardous chemical accidents [16]. Genetic fuzzy C-means were used for hazardous material transport based on clustering algorithm [17].

In this study, a new system was designed to measure the electromagnetic wave reflection coefficients of liquids. Hazardous liquids were separated from everyday liquids using the data obtained from the designed measurement system and hierarchical clustering. Although hierarchical clustering was used in the literature, the effect of linkage methods and distance metrics on clustering was not examined. In this study, different from other studies, a new measurement system was developed and the effects of different linkage methods and different distance metrics on hierarchical clustering were examined.

2. Materials and Methods

The antenna shown in Fig. 1 is located on a ground plane of a 10 cm x 10 cm square layer with a dielectric constant of 4.4. The thickness of the square layer is 1.6 mm. On this layer there is a layer made of copper with a circular geometry. The diameter of this circular sheet is 5.5 cm. Coaxial probe feeding method is used for feeding the antenna. The coaxial probe's inner conductor is connected to the antenna's radiation patch and the coaxial probe's outer conductor is connected to the antenna's ground plane. Antenna is feed by 50 Ohm SMA (SubMiniature version A) feed probe. The diameter of the antenna is calculated using Equation (1) and (2).

$$F = \frac{8,791 \times 10^9}{f_r \sqrt{\epsilon_r}} \quad (1)$$

$$a = \frac{F}{\left\{ 1 + \frac{2h}{\pi \epsilon_r F \left[\ln \left(\frac{\pi F}{2h} \right) + 1,7726 \right]^{1/2}} \right\}} \quad (2)$$

where ϵ_r represents relative permittivity of the substrate, f_r represents resonant frequency, h represents height of the substrate, and a represents radius of the patch, respectively.

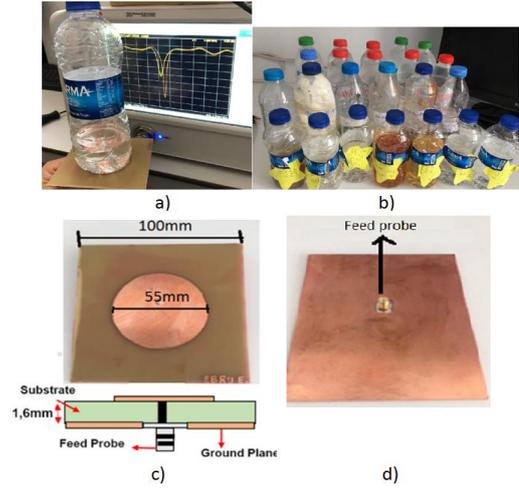


Figure 1. a) Measurement system used in this study b) Samples c) The geometry and front view of the patch antenna d) Back view of the patch antenna.

Cluster analysis, i.e. clustering, is a data mining technique that groups data in a given database according to its characteristics. Clustering analysis is an unsupervised learning method and does not include any training phase. In other words, only observations with input variables are given to the system. The target is not output. The system performs its learning alone without using a controller and tries to find the relationship between input variables by evaluating a certain affinity criterion. There are some distance metrics used in the calculations of the proximity criterion of clustering techniques. Commonly used ones of these distance metrics are Chebyshev, Euclidean, Manhattan and Minkowski.

2.1. Hierarchical Clustering Approaches and Methods

Hierarchical clustering has two main approaches, namely agglomerative and divisive [18]. In the agglomerative hierarchical clustering approach, each sample is taken as a cluster. By calculating the proximity of these clusters to each other, the closest clusters are combined. This process is continued until there is only one cluster. On the other hand, in the divisive hierarchical clustering approach, all samples are taken as a single set and subsets are created by making proximity calculations.

Hierarchical clustering methods used in this study are:
Single-linkage: It is based on the creation of clusters based on the smallest distance between cluster elements. This method is also known as the closest neighbor method. Mathematically, the distance between P_1 and P_2 sets $d(P_1, P_2)$ is expressed as in Equation (3).

$$d(P_1, P_2) = \text{Min}_{x_1 \in P_1, x_2 \in P_2} d(x_1, x_2) \quad (3)$$

where P_1 and P_2 represent any two sets of elements considered as clusters, and $d(x_1, x_2)$ represents the distance between the two elements x and y .

Complete linkage: It is based on the minimum value of the greatest distance between the elements of two sets. This distance is expressed as in Equation (4).

$$d(P_1, P_2) = \text{Max}_{x_1 \in P_1, x_2 \in P_2} d(x_1, x_2) \quad (4)$$

Average-linkage: It is based on the smallest value of the average distance between clusters. Its mathematical representation is given in Equation (5).

$$d(P_1, P_2) = \frac{1}{|P_1|} \frac{1}{|P_2|} \sum_{x_1 \in P_1} \sum_{x_2 \in P_2} d(x_1, x_2) \quad (5)$$

Centroid-linkage: It takes into account the distance between the two cluster centers. d represents the distance between two sets and is given in Equation (6).

$$d(P_1, P_2) = d\left(\left(\frac{1}{|P_1|} \sum_{x_1 \in P_1} \bar{x}_1\right), \left(\frac{1}{|P_2|} \sum_{x_2 \in P_2} \bar{x}_2\right)\right) \quad (6)$$

where P_1 represents the first cluster's center and P_2 represents the second cluster's center.

$$Td_{P_1 \cup P_2} = \sum_{x \in P_1 \cup P_2} d(x, \mu_{P_1 \cup P_2})^2 \quad (7)$$

Ward-linkage: It is a method that takes into account the average distance of the observations falling in the middle of a cluster from the observations in the same cluster, in other words, it is a method that aims to minimize the total within-cluster variance and makes use of total squares of deviation [19]. It is calculated using Equation (7).

3. Results and Discussion

In this study, S_{11} parameter measurements of 36 liquids, 7 hazardous and 29 non-hazardous liquids, were made and these data were used as input to the clustering algorithm. Using Weka platform, the liquids were divided into two clusters. The names of the liquids measured are given in Table 1 and measurement charts made are given in Fig. 2.

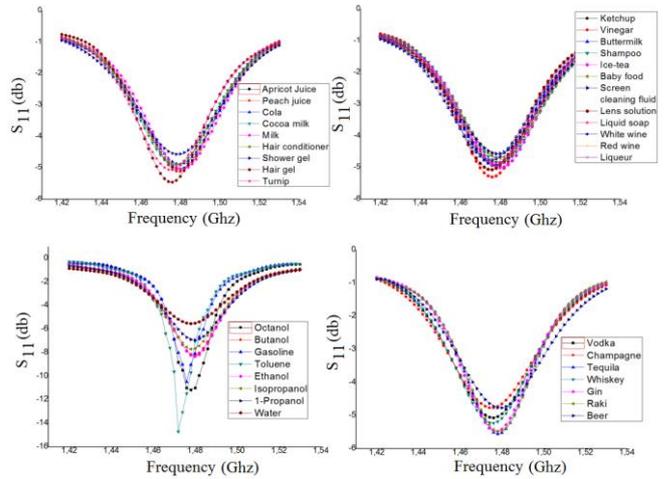


Figure 2. S_{11} measurements of the liquids used in this study.

Five different linkage methods and four different distance metrics were used to classify hazardous liquids. For this, a total of 20 clustering operations were performed using four different distance metrics for each linkage method and the results were examined. Liquids classified incorrectly are listed in Table 2.

Table 1. Liquids used in this study

Non-hazardous liquids	Non-hazardous liquids	Hazardous liquids
1 Apricot juice	16 Lens solution	30 Octanol
2 Peach juice	17 Liquid soap	31 Butanol
3 Cola	18 Water	32 Gasoline
4 Milk	19 White wine	33 Toluene
5 Hair conditioner	20 Red wine	34 Ethanol
6 Shower gel	21 Liqueur	35 1-propanol
7 Hair gel	22 Vodka	36 Isopropanol
8 Turnip juice	23 Champagne	
9 Ketchup	24 Shampoo	
10 Vinegar	25 Tequila	
11 Buttermilk	26 Whiskey	
12 Shampoo	27 Gin	
13 Ice-tea	28 Raki	
14 Baby food	29 Beer	
15 Screen cleaning fluid		

Table 2. Incorrectly classified liquids by different hierarchical clustering algorithms

Links	Chebyshev	Euclidean	Manhattan	Minkowski
Single	Octanol	Butanol	Butanol	Butanol
	Butanol	Ethanol	Ethanol	Ethanol
	Gasoline	1-	1-	1-
	Ethanol	Propanol	Propanol	Propanol
	1-Propanol	Isopropanol	Isopropanol	Isopropanol
Complete	Propanol	Isopropanol	Isopropanol	Isopropanol
	Isopropanol	Isopropanol	Isopropanol	Isopropanol
	Butanol	Butanol	Butanol	Butanol
	Ethanol	Ethanol	Ethanol	Ethanol
	1-Propanol	1-Propanol	1-Propanol	1-Propanol

	Toluene			
Average	Octanol	Butanol	Butanol	Butanol
	Butanol	Ethanol	Ethanol	Ethanol
	Gasoline	1-	1-	1-
	Ethanol	Propanol	Propanol	Propanol
	1-Propanol	Isopropanol	Isopropanol	Isopropanol
Centroid	Octanol	Butanol	Butanol	Butanol
	Butanol	Ethanol	Ethanol	Ethanol
	Gasoline	1-	1-	1-
	Ethanol	Propanol	Propanol	Propanol
	1-Propanol	Isopropanol	Isopropanol	Isopropanol
Ward	-----	-----	-----	-----

As the results in Table 2 proves, the worst results were obtained when Chebyshev was used as the distance metric. When it was used, Butanol, Ethanol, 1-Propanol, Isopropanol, Gasoline and Octanol were misclassified and only Toluene was classified correctly among hazardous liquids. On the other hand, when Ward linkage method was used, all the liquids were correctly classified regardless of the employed distance metric. Accuracy was used as the performance metric to numerically compare the performance of the linkage methods and distance metrics. It is the ratio of correct predictions to the total number of predictions. In Fig. 3, the accuracy values of the linkage methods for different distance metrics are shown. As can be seen in Fig. 3, Ward linkage method was the most successful method with 100% accuracy regardless of the employed distance metric.

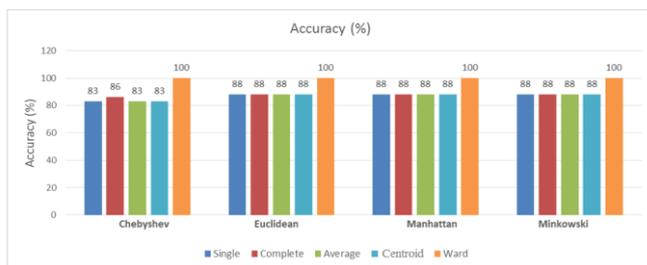


Figure 3. Accuracy rates of the hierarchical clustering methods used in the experiments.

In order to examine the effect of linkage methods and distance metrics on the clustering result, a number was assigned to each liquid and clustering was performed. The most accurate result in hazardous liquid detection was obtained when Ward linkage method was used. The results obtained using clustering based on Ward linkage method are shown in Figure 4.

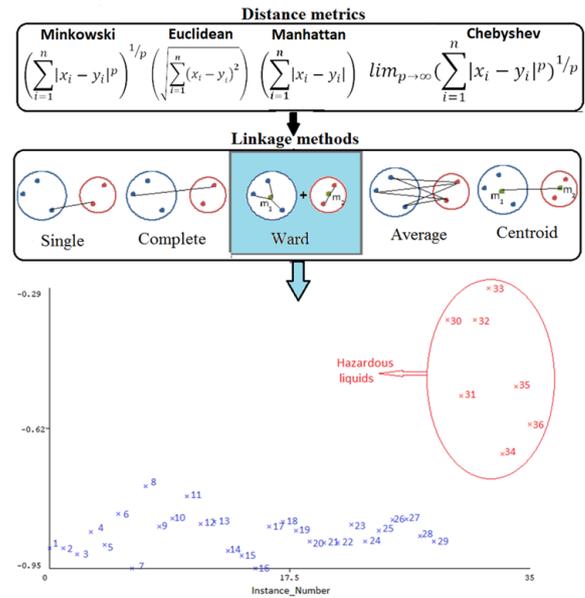


Figure 4. Results of the clustering made using Ward linkage method.

Fig. 5 is a group of confusion matrices that show the relation between the number of actual Hazardous (H) and Non-Hazardous (NH) liquids and the predicted ones when different linkage methods and distance metrics were used. The boxes in red in the matrix indicate the number of liquids that the algorithm made incorrect predictions. For instance, in the classification made using Chebyshev distance and Single linkage method, 6 out of 7 hazardous liquids were classified incorrectly and 1 was correctly classified, and all non-hazardous liquids were classified correctly. As it can be seen in the figure, when Ward linkage method was used, all the liquids were correctly classified regardless of the distance metrics.

		Chebyshev Predicted		Euclidean Predicted		Manhattan Predicted		Minkowski Predicted		
Single	Actual	H	NH	H	NH	H	NH	H	NH	
		1	6	3	4	3	4	3	4	
Single	Actual	NH	0	29	0	29	0	29	0	29
Complete	Actual	H	NH	H	NH	H	NH	H	NH	
		2	5	3	4	3	4	3	4	
Complete	Actual	NH	0	29	0	29	0	29	0	29
Average	Actual	H	NH	H	NH	H	NH	H	NH	
		1	6	3	4	3	4	3	4	
Average	Actual	NH	0	29	0	29	0	29	0	29
Centroid	Actual	H	NH	H	NH	H	NH	H	NH	
		1	6	3	4	3	4	3	4	
Centroid	Actual	NH	0	29	0	29	0	29	0	29
Ward	Actual	H	NH	H	NH	H	NH	H	NH	
		7	0	7	0	7	0	7	0	
Ward	Actual	NH	0	29	0	29	0	29	0	29

Figure 5. Confusion matrices for the linkage methods and distance metrics.

4. Conclusion

Since detection of hazardous liquids used in explosive production is critical and existing methods are either not practical or have some shortcomings, there is an imperative need for practical and reliable methods for the detection of hazardous liquids. In this study, a circular patch antenna for this purpose was designed and connected to a vector network analyzer. Then, a group of S_{11} measurements of 36 liquids in 1.42-1.53 GHz was made. Finally, five different hierarchical clustering algorithms was used to classify sample liquids as hazardous or not. As proven in the experiments realized using different distance metrics, Ward linkage method achieved 100% in the detection of hazardous liquids.

5. References

- [1] Zandieh, O. and Kim, S., "Sensitive and selective detection of adsorbed explosive molecules using opto-calorimetric infrared spectroscopy and micro-differential thermal analysis," *Sensors and Actuators B: Chemical*, vol. 231, pp. 393-398, 2016.
- [2] Gares, K., L. Hufziger, K. T., Bykov, S. V. and Asher, S. A., "Review of explosive detection methodologies and the emergence of standoff deep UV resonance Raman," *Journal of Raman Spectroscopy*, vol. 47, no. 1, pp. 124-141, 2016.
- [3] Zarei, A. R. and Ghazanchayi, B., "Design and fabrication of optical chemical sensor for detection of nitroaromatic explosives based on fluorescence quenching of phenol red immobilized poly (vinyl alcohol) membrane," *Talanta*, vol. 150, pp. 162-168, 2016.
- [4] Mäkinen, M., Nousiainen, M. and Sillanpää, M., "Ion spectrometric detection technologies for ultra-traces of explosives: A review," *Mass spectrometry reviews*, vol. 30, no. 5, pp. 940-973, 2011.
- [5] Chrzanowski, L. von, Beckmann, J. Marchetti, B., Ewert, U. and Schade, U., "Terahertz time domain spectroscopy for non-destructive testing of hazardous liquids," *Materials Testing*, vol. 54, no. 6, pp. 444-450, 2012.
- [6] Botti, S., Almaviva, S., Cantarini, L., Palucci, A., Puiu, A. and Rufoloni, A., "Trace level detection and identification of nitro-based explosives by surface-enhanced Raman spectroscopy," *Journal of Raman Spectroscopy*, vol. 44, no. 3, pp. 463-468, 2013.
- [7] Ramírez-Cedeño, M. L., Ortiz-Rivera, W., Pacheco-Londoño, L. C. and Hernández-Rivera, S. P., "Remote detection of hazardous liquids concealed in glass and plastic containers," *IEEE Sensors Journal*, vol. 10, no. 3, pp. 693-698, 2010.
- [8] Krylatykh, N. A., Yakh'ya, V. F., Fakhrutdinov, A. R., Anashkin, V. N., Shagalov, V. A. and Khabipov, R. S., "Detection of explosive precursors using low-field magnetic resonance imaging," *Applied Magnetic Resonance*, vol. 47, no. 8, pp. 915-924, 2016.
- [9] Castro-Suarez, J. R., Pacheco-Londoño, L. C., Aparicio-Bolaño, J. and Hernández-Rivera, S. P., "Active mode remote infrared spectroscopy detection of TNT and PETN on aluminum substrates," *Journal of Spectroscopy*, vol. 2017, 2017.
- [10] Kehres, J., Lyksborg, M. and Olsen, U. L., "Threat detection of liquid explosives and precursors from their x-ray scattering pattern using energy dispersive detector technology," in *Radiation Detectors in Medicine, Industry, and National Security XVIII*, 2017, vol. 10393: International Society for Optics and Photonics, p. 1039302.
- [11] Singh, P., Bhamidipati, S., Singh, R., Smith, R. and Nelson, P., "Evaluation of in-line sensors for prediction of soluble and total solids/moisture in continuous processing of fruit juices," *Food Control*, vol. 7, no. 3, pp. 141-148, 1996.
- [12] Abduljabar, A. A., Hamzah, H. and Porch, A., "Multi-resonators, microwave microfluidic sensor for liquid characterization," *Microwave and Optical Technology Letters*, vol. 63, no. 4, pp. 1042-1047, 2021.
- [13] Hafdi, Z., Tao, J. and Chaabi, A., "Microstrip coupled high sensitivity sensor for water ethanol mixture characterization," *Frequenz*, vol. 75, no. 1-2, pp. 1-7, 2021.
- [14] Efeoğlu E. and Tuna, G., "The Use of Microwave and K* Algorithm in Determination of Alcohol Concentration in Liquids," *Russian Journal of Nondestructive Testing*, vol. 56, no. 8, pp. 689-697, 2020.
- [15] Yuan, S., Zhang, Z., Sun, Y., Kwon, J. S.-I. and Mashuga, C. V., "Liquid flammability ratings predicted by machine learning considering aerosolization," *Journal of Hazardous Materials*, vol. 386, p. 121640, 2020.
- [16] Deng, F., Gu, W., Zeng, W., Zhang, Z. and Wang, F., "Hazardous Chemical Accident Prevention Based on K-Means Clustering Analysis of Incident Information," *IEEE Access*, vol. 8, pp. 180171-180183, 2020.
- [17] Wang, X. and Wang, H., "Driving behavior clustering for hazardous material transportation based on genetic fuzzy C-means algorithm," *IEEE Access*, vol. 8, pp. 11289-11296, 2020.
- [18] Rafsanjani, M. K., Varzaneh, Z. A. and Chukanlo, N. E., "A survey of hierarchical clustering algorithms," *The Journal of Mathematics and Computer Science*, vol. 5, no. 3, pp. 229-240, 2012.
- [19] Murtagh, F. and Contreras, P., "Algorithms for hierarchical clustering: an overview, II," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 7, no. 6, p. e1219, 2017.