RADIO WAVE METHODS

The Use of Microwave and K^* Algorithm in Determination of Alcohol Concentration in Liquids

Ebru Efeoğlu^a and Gurkan Tuna^{a, *}

^aTrakya University, Edirne, Turkey

*e-mail: gurkan@trakya.edu.tr

Received May 15, 2020; revised June 11, 2020; accepted June 11, 2020

Abstract—Alcohol is widely used in various fields today. Alcohol overdose and its improper use can cause many health problems. Drinking liquids with high alcohol concentration, especially methanol, causes poisoning, and the use of colognes with high alcohol concentration causes various skin conditions and respiratory diseases. Therefore, for public health it is important to determine the type and concentration of alcohol in beverages and liquid products. In this paper, K^* algorithm is used to detect and classify liquids containing high concentration of alcohol by measuring the scattering parameter of the liquids in the microwave frequency band. For this purpose, scatter parameter values of aqueous solutions of ethanol, methanol, 1-propanol and 2-propanol at different concentrations have been used as training sets. Commercial liquids with known ethanol concentrations and liquids with other known alcohol concentrations have been tested to prove the accuracy of the proposed approach. As the results show the proposed approach can classify liquids containing alcohol and their alcohol concentrations successfully and with high accuracy.

Keywords: classification, K^* algorithm, microwaves, patch antenna, vector network analyzer

DOI: 10.1134/S1061830920080033

INTRODUCTION

Pollution resulting from industrial processes and increasing population is a major threat to human health, plants and animals [1]. Methanol, which is one of the harmful liquids, is widely used in laboratories and industry as a synthetic precursor and solvent. It is a highly toxic substance. The smell of this liquid and its contact with the skin harms human health. Recently, hand disinfectant and cologne have been used frequently to stop the spread of epidemics such as corona virus. Alcohol-based hand disinfectants mostly include ethanol, isopropyl alcohol, n-propyl alcohol, or combinations thereof [2]. When used improperly, these can be toxic to humans. For instance, if accidentally or intentionally swallowed, they can cause central nervous system, respiratory diseases and eye diseases [3]. Hand disinfectants must not contain methanol [4] and the presence of methanol in the disinfectants found in the market poses a serious threat to public health. Similarly, high alcohol content in colognes may cause skin and health problems. Moreover, consumption of illegally produced alcohol is an important public health problem in many parts of the world [5, 6].

Alcohol content may be higher than it is expected to be and harmful to health in the alcoholic drinks which are produced illegally and without quality control processes during production. In addition, methanol can be added to alcoholic drinks by the manufacturers to increase the amount of alcohol and reduce the cost. Because methanol is easier to obtain and is much cheaper than ethanol. The final product of ethanol is acetic acid, which can be used in our body and is used within the framework of acid-base balance. On the other hand, the final product of methanol is formic acid, which is not an acid that can be used in our body. Formic acid is slowly excreted in the urine and through the respiratory tract. When it accumulates in the blood, it causes acidosis. For this reason, when illegally produced beverages containing methanol are consumed, vision problems or death might occur due to formaldehyde and formic acid accumulated in the blood by the destruction of methanol. Since beverages containing high concentrations of methanol can cause serious health problems or deaths when consumed [7] and toxic alcohol consumption is an important cause of death and diseases worldwide [8], the presence of methanol in beverages must not be allowed [9].

Fig. 1. Experimental setup used for liquid classification.

The effects of these chemicals on the environment are seen as drinking water, air and sea pollution. In order to eliminate or minimize those effects, different water treatment approaches are used [10]. Because, liquids like methanol, ethanol, 2-propanol are toxic not only for humans, but also for aquatic organisms [11]. Marine pollution and lake pollution cause oxygen depletion in water and as a result can threaten underwater life. Therefore, the Eco toxicity assessment of chemicals can play a key role in deriving threshold values in terms of safety and acceptable concentrations, but the reliability and suitability of Eco toxicological data need be considered [12].

Microwave-based measurement methods are rapid, low-cost and non-hazardous, and unaffected by environmental conditions [13, 14]. They are generally used to measure the relative permeability values of materials. For this purpose, a formula model optimized with artificial bee colony (ABC) algorithm has been proposed to calculate the relative permeability of materials using the microwave method [15]. A number of microwave techniques have been reported in the literature and have been used in practice for liquids. They can be categorized into two main groups: open-ended coaxial probe [14] and the Free space [16]. Each group has its own advantages, disadvantages and application types. For example, the coaxial probe technique is prone to errors caused by improper contact between the probe and the material surface (e.g. air gaps or air bubbles). Time-domain reflectometers are expensive [17]. On the other hand, particular attention should be paid to the horn antenna selection, the design of the sample holder and the sample geometry and location for the free space technique [18].

Classification and clustering techniques are commonly used in environmental applications. Total heavy metal concentrations of sewage sludge obtained from mechanical-biological wastewater treatment plants were characterized by hierarchical cluster analysis [19]. Also, cluster analysis was used to investigate elemental concentrations of river sediments [20]. For the detection and determination of nickel in solution, a quartz crystal nanobalance and principal component analysis were used [21]. Along with the principal component analysis, a selective determination of ethyl acetate, acetone, ethanol and methyl ethyl ketone was realized using the quartz crystal nanobalance [22]. Ethanol and methanol concentrations were studied using ultrasonic velocity measurements in ethanol–water and methanol–water mixtures in solutions obtained with alcohol water mixture [23]. Moreover, aqueous alcohol solutions were characterized in bottles by THz spectroscopy method [24].

EXPERIMENTAL SETUP, METHODOLOGY AND RESULTS

Scattering parameters or S parameters (elements of the scattering matrix or S-matrix) describe the electrical behavior of linear electrical networks when exposed to various steady-state stimuli by electrical signals. S_{11} is the reflection coefficient and S_{21} is the transmission coefficient. In the study, S_{11} measurements of the liquids were made. As shown in Fig. 1, the measurement device consists of a vector network analyzer (VNA) and a circular patch antenna. Liquid measurements can be carried out in closed containers, as well as by dipping into the liquid for containers larger than 10 cm. The resonance frequency of the antenna fed with a 50 Ohm SMA (SubMiniature version A) feed probe is 1.5 GHz. Figure 2 shows the schematic design of the network analyzer and antenna.

Fig. 2. Design of the vector network analyzer and antenna.

Some of the signal from the source is conveyed while some of it is reflected. The reflection coefficient is defined as the ratio of the amplitude of the reflected signal to the amplitude of the transmitted signal, and is calculated using (1).

$$\text{Reflection coefficient} = \text{Reflected/Incident (A/R)}. \quad (1)$$

Two different experimental studies were performed in this study. In the first one, it was determined whether a measured liquid contains alcohol or not and the name of the alcohol if it contains alcohol. In the second one, both the type and the alcohol concentration of the liquid was determined. The flow of the approach is shown in Fig. 3.

K^* Algorithm

K^* algorithm uses the entropic distance measure to determine the spectral distance or similarity between the data [25]. This distance is based on information theory. In information theory, entropy is a measure of uncertainty about random variables [26], the estimate of the distance between data is the complexity in the conversion of one data into another. This complexity is calculated in two steps. In the first step, a finite number of sets of transformations are defined, which are defined as feature-to-feature conversions. In the second step, the transformation of one feature (a) to another feature (b) is defined as a series of limited transformations that begin at (a) and end at (b) [25]. K^* algorithm calculates the shortest distance that connects the two features using the Kolmogorov distance. In this case, K^* distance is the sum of all possible transformations between the two properties [27].

Probability function P^* , as the probability of all paths from property a to property b with t transforms,

$$P^*(b/a) = \sum_{\bar{t} \in P: \bar{t}(a)=b} p(\bar{t}). \quad (2)$$

Then K^* function becomes

$$K^*(b/a) = -\log_2 P^*(b/a). \quad (3)$$

K^* function can be used for numerical or symbolic data sets. In the use of the function for numerical class values, the parameter x_0 is determined by the user prior to training. In both cases, the number of samples included in the probability distribution varies from the nearest neighborhood distribution value 1 to the N value, where all samples are of equal weight [25]. Effective number of samples for any P^* function is

$$n_0 \leq \frac{\left(\sum_b P^*(b/a) \right)^2}{\sum_b P^*(b/a)^2} \leq N. \quad (4)$$

Fig. 3. Approach used to classify liquid type and alcohol concentration.

Here, N indicates the number of all training samples, and n_0 indicates the number of training samples from the training sample at the shortest distance. K^* algorithm determines the value of a for parameters x_0 and s by selecting a value between n_0 and N . Thus, the selected no gives the closest neighborhood algorithm and the number of selected N equal weight samples. The suitability of these determined numbers is determined using the blending parameter (b). The blending parameter ranges from $b = 0\%$ (for n_0) $b = 100\%$ (for N).

Performance Metrics

There are some performance metrics used in performance evaluation of recognition and classification algorithms. The most popular of these are accuracy, root mean squared error (RMS) and Kappa statistics. Accuracy gives the number of samples correctly classified from all samples (5).

$$\text{Accuracy} = \frac{\text{True positive} + \text{True negative}}{\text{Number of instances}}. \quad (5)$$

RMS is used to scale the difference between actual and predicted values. It is calculated as given in (6).

$$\text{RMS} = \sqrt{\frac{(P_1 - a_1)^2 + \dots + (P_n - P_n)^2}{n}}, \quad (6)$$

where P represents the estimated values and a represents the actual values. As RMS value approaches zero, the accuracy of the predictive model increases, and the accuracy of the predictive model decreases as RMS value moves away from zero. Therefore, it is desired that RMS value be low, preferably very close to zero.

Kappa value is used to measure the consistency between the predicted and observed classifications on a group of data. Kappa value is calculated using (7).

$$K = \frac{P(a) - P(e)}{1 - P(e)}, \quad (7)$$

Fig. 4. S_{11} parameters of the liquids used in the experiment.

where $P(a)$ indicates the accuracy of the algorithm, and $P(e)$ is the weighted average of the expected accuracy of the algorithm, which makes random predictions in the same dataset. The closer the value is to 1, the greater the fit.

RESULTS

In the first experiment that aimed to determine whether a liquid contains alcohol or not, 14 different liquids were used. 10 of the liquids used are the drinks we use in our daily lives and liquid that touches human skin like hand soap, shower gel, etc. Four different types of alcohol were used. The names of the liquids used in the measurements are given in Table 1.

S_{11} parameters of the liquids are shown in Fig. 4. In this figure, the horizontal axis shows the amplitude of the S_{11} parameters and the vertical axis shows the frequency in GHz. As shown in the Fig. 4, the liquid with the highest amplitude value is 1-propanol, and the liquid with the least amplitude value is methanol. As seen in Fig. 5, non-alcoholic drinks were also grouped in a separate class.

In the second experiment, 45 different liquids were used and both the types of the liquids and the alcohol concentration of the liquids were predicted. Aqueous solutions of different concentrations of ethanol, methanol, 1-propanol and 2-propanol alcohols (10, 20, ..., 100% by volume) were prepared and S_{11} parameters of liquids of these liquid solutions were used as training data. The accuracy of the algorithm was then tested using trademark alcoholic beverages (red wine, whiskey, vodka and white wine) and liquids of different alcohol concentrations. S_{11} parameters of sample alcohol-water solutions with different concentrations of alcohol are shown in Fig. 6. Accuracy of the proposed approach for those different liquids is presented in Table 2. As can be seen, except for the 40% 2-propanol case, the accuracy of the proposed approach is quite high, very close to 100%.

Overall performance of the proposed approach is shown in Fig. 7. As can be seen, the algorithm correctly classified all 14 liquids (in blue) and 45 of 46 liquids (red one). It has 100% accuracy for classifying the liquids and 97% accuracy for classifying the alcohol type and concentration. In addition, Kappa value

Table 1. Liquid types used in this experiment

Liquid types (No-alcohol)		Alcohols
cherry juice	shower gel	ethanol
cola	liquid soap	methanol
milk	shampoo	1-propanol
buttermilk	body lotion	2-propanol
water	ice-tea

Fig. 5. Classification of liquids by liquid type.

Fig. 6. S_{11} parameters of the water-alcohol solutions of different alcohol concentrations (a) ethanol, (b) methanol, (c) 1-propanol, (d) 2-propanol.

Fig. 7. Performance metrics of the algorithm.

is 1 for classifying the liquid type and 0.977 for classifying the alcohol type and concentration. In both of them, RMS value is quite close to zero, 0.001 and 0.023, respectively.

CONCLUSIONS

Preventing the high concentrations of alcohols and their improper uses is important to minimize and eliminate risks that threaten human health. In this study, an antenna was designed and manufactured for use in the measurement of reflection coefficients that were used in the classification phase. S_{11} parameters

Table 2. Accuracy of the proposed approach for different liquids

Alcohol type	Actual, %	Predicted, %	True/false
Red wine	12% ethanol	10% ethanol	true
Vodka	40% ethanol	40% ethanol	true
Whiskey	40% ethanol	40% ethanol	true
White wine	12% ethanol	10% ethanol	true
Ethanol	(10, 20, 30, 40, 50, 60, 70, 80, 90, 100)% ethanol	(10, 20, 30, 40, 50, 60, 70, 80, 90, 100)% ethanol	true
Methanol	(10, 20, 30, 40, 50, 60, 70, 80, 90, 100)% methanol	(10, 20, 30, 40, 50, 60, 70, 80, 90, 100)% methanol	true
1-Propanol	(10, 20, 30, 40, 50, 60, 70, 80, 90, 100)% 1-propanol	(10, 20, 30, 40, 50, 60, 70, 80, 90, 100)% 1-propanol	true
2-Propanol	(10, 20, 30, 50, 60, 70, 80, 90, 100)% 2-propanol	(10, 20, 30, 50, 60, 70, 80, 90, 100)% 2-propanol	true
2-Propanol	40% 2-propanol	100% methanol	false

obtained with the antenna connected to the vector network analyzer were processed using the K^* algorithm-based classification phases. As the results of the classification phases show the proposed approach can classify liquids containing alcohol and their alcohol concentrations successfully and with high accuracy. A prototype system based on the proposed approach and that is lightweight and can be practically used is under development.

REFERENCES

1. Valavanidis, A., Fiotakis, K., and Vlachogianni, T., Airborne particulate matter and human health: Toxicological assessment and importance of size and composition of particles for oxidative damage and carcinogenic mechanisms, *J. Environ. Sci. Health, Part C*, 2008, vol. 26, no. 4, pp. 339–362.
2. Pires, D., de Kraker, M.E.A., Tartari, E., Abbas, M., and Pittet, D., ‘Fight antibiotic resistance—It’s in your hands’: Call from the World Health Organization for 5th May 2017, *Clin. Infect. Dis.*, 2017, vol. 64, no. 12, pp. 1780–1783.
3. Slaughter, R., Mason, R., Beasley, D., Vale, J., and Schep, L., Isopropanol poisoning, *Clin. Toxicol.*, 2014, vol. 52, no. 5, pp. 470–478.
4. Chan, A.P. and Chan, T.Y., Methanol as an unlisted ingredient in supposedly alcohol-based hand rub can pose serious health risk, *Int. J. Environ. Res. Public Health*, 2018, vol. 15, no. 7, p. 1440.
5. Lachenmeier, D.W., Rehm, J., and Gmel, G., Surrogate alcohol: What do we know and where do we go? *Alcohol.: Clin. Exp. Res.*, 2007, vol. 31, no. 10, pp. 1613–1624.
6. Rehm, J., Kailasapillai, S., Larsen, E., Rehm, M.X., Samokhvalov, A.V., Shield, K.D., Roerecke, M., and Lachenmeier, D.W., A systematic review of the epidemiology of unrecorded alcohol consumption and the chemical composition of unrecorded alcohol, *Addiction*, 2014, vol. 109, no. 6, pp. 880–893.
7. Destanoğlu, O. and İsmail, A., Determination and evaluation of methanol, ethanol and higher alcohols in legally and illegally produced alcoholic beverages, *J. Turk. Chem. Soc., Sec. A: Chem.*, 2019, vol. 6, no. 1, pp. 21–28.
8. Organization, W. H., *Global Status Report on Alcohol and Health 2018*, W.H.O., 2019.
9. Zamani, N., Rafizadeh, A., Hassanian-Moghaddam, H., Akhavan-Tavakoli, A., Ghorbani-Samin, M., Akhgari, M., and Shariati, S., Evaluation of methanol content of illegal beverages using GC and an easier modified Chromotropic acid method; a cross sectional study, *Subst. Abuse. Treat. Prev. Policy*, 2019, vol. 14, no. 1, p. 56.
10. Chowdhary, P., Raj, A., and Bharagava, R.N., Environmental pollution and health hazards from distillery wastewater and treatment approaches to combat the environmental threats: A review, *Chemosphere*, 2018, vol. 194, pp. 229–246.
11. Network, D.R. and Bishop, R.E., Chemical and Biological Hazards Posed by Drilling Exploratory Shale Gas Wells in Pennsylvania’s Delaware River Basin Report for the Delaware River Basin Commission Exploratory Well Hearing, 2010.
12. Moermond, C.T., Kase, R., Korkaric, M., and Ågerstrand, M., CRED: Criteria for reporting and evaluating ecotoxicity data, *Environ. Toxicol. Chem.*, 2016, vol. 35, no. 5, pp. 1297–1309.
13. Kharkovsky, S. and Zoughi, R., Microwave and millimeter wave nondestructive testing and evaluation—Overview and recent advances, *IEEE Instrum. Meas. Mag.*, 2007, vol. 10, no. 2, pp. 26–38.
14. Li, Z., Haigh, A., Soutis, C., Gibson, A., and Sloan, R., Microwaves sensor for wind turbine blade inspection, *Appl. Compos. Mater.*, 2017, vol. 24, no. 2, pp. 495–512.
15. Tekbas, M., Toktas, A., and Ustun, D., A formulaic model calculating the permittivity of testing materials placed on a circular patch antenna, in *2019 XXIVth Int. Semin./Workshop Direct Inverse Probl. Electromagn. Acoust. Wave Theory (DIPED)*, 2019, IEEE, 2019, pp. 88–92.
16. Jose, K., Varadan, V., and Varadan, V., Wideband and noncontact characterization of the complex permittivity of liquids, *Microwave Opt. Technol. Lett.*, 2001, vol. 30, no. 2, pp. 75–79.
17. Venkatesh, M. and Raghavan, G., An overview of dielectric properties measuring techniques, *Can. Biosyst. Eng.*, 2005, vol. 47, no. 7, pp. 15–30.
18. Li, Z., Haigh, A., Soutis, C., Gibson, A., and Sloan, R., Evaluation of water content in honey using microwave transmission line technique, *J. Food Eng.*, 2017, vol. 215, pp. 113–125.
19. Gusiatin, Z.M., Kulikowska, D., Klik, B.K., and Hajdukiewicz, K., Ecological risk assessment of sewage sludge from municipal wastewater treatment plants: A case study, *J. Environ. Sci. Health, Part A*, 2018, vol. 53, no. 13, pp. 1167–1176.
20. Ewa, I., Oladipo, O., and Dim, L., Cluster analysis of elemental concentrations of cored Nigerian river sediments, *J. Environ. Sci. & Health, Part A*, 1992, vol. 27, no. 1, pp. 1–11.

21. Mirmohseni, A., Shojaei, M., Feizi, M.A.H., Azhar, F.F., and Rastgouye-Houjaghan, M., Application of quartz crystal nanobalance and principal component analysis for detection and determination of nickel in solution, *J. Environ. Sci. & Health, Part A*, 2010, vol. 45, no. 9, pp. 1119–1125.
22. Mirmohseni, A., Razzaghi, M., Pourata, R., Rastgouye-Hojagan, M., and Zavareh, S., Selective determination of ethyl acetate, acetone, ethanol, and methyl ethyl ketone using quartz crystal nanobalance combined with principle component analysis, *J. Environ. Sci. & Health, Part A*, 2009, vol. 44, no. 9, pp. 847–853.
23. Vatandas, M., Koc, A.B., and Koc, C., Ultrasonic velocity measurements in ethanol–water and methanol–water mixtures, *Eur. Food Res. Technol.*, 2007, vol. 225, nos. 3–4, pp. 525–532.
24. Jepsen, P.U., Jensen, J.K., and Møller, U., Characterization of aqueous alcohol solutions in bottles with THz reflection spectroscopy, *Opt. Express*, 2008, vol. 16, no. 13, pp. 9318–9331.
25. Cleary, J.G. and Trigg, L.E., K^* : An instance-based learner using an entropic distance measure, in *Machine Learning Proceedings 1995*, Elsevier, 1995, pp. 108–114.
26. Shannon, C. E., A mathematical theory of communication, *Bell Syst. Tech. J.*, 1948, vol. 27, no. 3, pp. 379–423.
27. Piramuthu, S. and Sikora, R.T., Iterative feature construction for improving inductive learning algorithms, *Expert Syst. Appl.*, 2009, vol. 36, no. 2, pp. 3401–3406.