On The Feasibility of Estimating Soluble Sugar Content Using Millimeter-wave

Zhicheng Yang
Dept. of Computer Science
UC Davis
Davis, CA, USA
zcyang@ucdavis.edu

Parth H. Pathak
Dept. of Computer Science
George Mason University
Fairfax, VA, USA
phpathak@gmu.edu

Mo Sha
Dept. of Computer Science
SUNY Binghamton
Binghamton, NY, USA
msha@binghamton.edu

Tingting Zhu
Dept. of Plant Sciences
UC Davis
Davis, CA, USA
tngzhu@ucdavis.edu

Junai Gan
Dept. of Food Sci and Tech
UC Davis
Davis, CA, USA
jagan@ucdavis.edu

Pengfei Hu
Dept. of Computer Science
UC Davis
Davis, CA, USA
pfhu@ucdavis.edu

Prasant Mohapatra
Dept. of Computer Science
UC Davis
Davis, CA, USA
pmohapatra@ucdavis.edu

ABSTRACT

With the development of novel sensing techniques, continuous monitoring, data-driven inferences, precision irrigation control, and intelligent Internet-of-Things (IoT) systems, agriculture sector is witnessing a revolution. Specialized devices based on infrared and laser are developed to assist farmers in assessing the produce quality, especially its sugar content. However, such devices are expensive and not readily available to consumers. In this paper, we investigate the feasibility of using 60 GHz millimeter-wave (mmWave) signal as a ubiquitous and non-invasive way to estimate the Soluble Sugar Content (SSC) in fruits. With the rapid development in the mmWave technology, 60 GHz WiFi is likely to become pervasive in future mobile devices. Our study shows that when 60 GHz WiFi signals reflect from a fruit, the reflection can be used to infer the fruit's sugar content. We identify the underlying reasons of variations in reflection signals with varying SSC and study the impact of size, shape and density of fruits on reflections. We then develop statistical features based on received signal strength and amplitude, and use them to design regression-based estimation models. With an extensive evaluation with 300 fruit samples, we find that our proposed technique can estimate SSC in three different type of fruits with an average correlation coefficient of 85%. Our prediction errors are within the range of user's taste perception.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous computing**; • **Applied computing** → *Consumer health*; Agriculture;

KEYWORDS

Millimeter-wave sensing; Soluble sugar content; Fruit quality assessment; Smart agriculture

1 INTRODUCTION

The rapid advancements in the Internet of Things (IoT) technologies have paved the way for smart agriculture. Recent research has addressed a variety of challenges toward this goal by designing intelligent IoT systems [50]. At the same time, significant efforts have been made to develop new approaches to monitor product quality. For instance, recent studies propose to use the Near Infrared (NIR) spectroscopy and laser imaging to assess fruit quality [40, 48] and the Soluble Sugar Content (SSC) of a fruit [18]. Such approaches can help farmers to carry out informed harvesting to maximize the yield. Estimating the quality of a fruit and predicting the precise time of consumption are also important for consumers. With such techniques, consumers can better choose the fruits based on their diet and taste preferences and better determine the precise time of consuming the fruit for highest satisfaction which in turn leads to better eating habits and healthy diet [12]. However, the existing approaches either rely on specialized, expensive devices or require invasive testing, making them infeasible to be used by consumers. Therefore, low-cost, non-invasive estimation of sugar content in fruits remains an open problem.

With the low-cost off-the-shelf devices being developed, 60 GHz mmWave WiFi devices are likely to become popular in future mobile and wearable devices. In this paper, we investigate the feasibility of using 60 GHz mmWave signal for non-invasive estimation of sugar content in fruits. Compared to microwave frequency band, mmWave can provide a better sensing resolution in the desired direction. We find that when 60 GHz WiFi signal reflects from a fruit, the received reflection can be used to infer the sugar content in the fruit. Specifically, we find that as the sugar content of a fruit increases, the Received Signal Strength (RSS) of the reflected signal decreases. Given that 60 GHz mmWave signals are transmitted using a directional antenna (to combat higher attenuation), a phased

antenna array on a smartphone can be used to transmit a mmWave signal probe to a fruit. The received signal can then be analyzed to estimate the sugar content.

Developing a mmWave system that estimates the sugar content in fruits poses unique challenges. First, it is not clear what are the underlying reasons of variation in RSS with changes in SSC. Understanding these reasons is essential to develop an accurate SSC estimation technique. Second, different fruits and their cultivar are different in size, shape, and density. These translate to factors such as fruit diameter, fruit surface roughness, flesh density and structure. It is crucial to characterize the impact of these factors on reflection in order to distill the SSC related impact. Third, depending on whether coarse-grained statistic like RSS or fine-grained raw signal is available after reflection, it is necessary to design a feature space that can be used by the machine learning algorithms to estimate SSC. Lastly, given the structural diversity in different fruit types and even the samples of the same fruit type, a large amount of reflection instances should be collected and analyzed for a robust evaluation of the claims. We address the above challenges in this paper. Specifically, the paper makes the following contributions:

(1) We show that it is feasible to use 60 GHz mmWave signal reflection to estimate sugar content. We use this insight to develop first-of-its-kind 60 GHz mmWave signal based SSC estimation technique for fruits. As we discussed before, the presented technique can be implemented on upcoming mobile devices (like smartphone and smartwatch) equipped with 60 GHz WiFi radio.

(2) We systematically study how varying levels of SSC affect the signal permittivity and in turn change the reflection. The relationships are studied through experimentation built on theoretical underpinning. We then investigate factors such as fruit diameter, fruit surface roughness, and flesh density and their impacts on the reflection. The studies are carried out for sugar solutions (for controlled experiments) and off-the-shelf fruits.

(3) We use our 60 GHz mmWave testbed to measure the reflections using coarse-grained RSS as well as fine-grained properties of the raw signal (e.g., amplitude). We identify statistical features that show a strong correlation with SSC and use them to develop regression models for estimation. We also develop estimation models that only utilize RSS given that most commodity radios can report RSS.

(4) We perform an extensive evaluation of our proposed technique using 300 different fruit samples (100 apples, 100 oranges and 100 kiwis), and find that our proposed approach can estimate the SSC with an average correlation coefficient of 85% (apples: 86%, oranges: 85% and kiwis: 84%) with four measurements per fruit and measurement time of as low as 1.5s per measurement. Our RSS-based estimation also achieves the mean correlation of 78% for regression models developed for each fruit type. For instance, users can taste the difference between apples when their Brix is different by more than 1 Bx [24]. Our prediction errors are within this range.

The remaining paper is organized as follows. Section 8 discusses related work. In Section 2, we first show the feasibility of estimating SSC using mmWave signal. Section 3 provides the underlying

(a) Measured Brix values of three simple sugar (sucrose, fructose and glucose) solutions

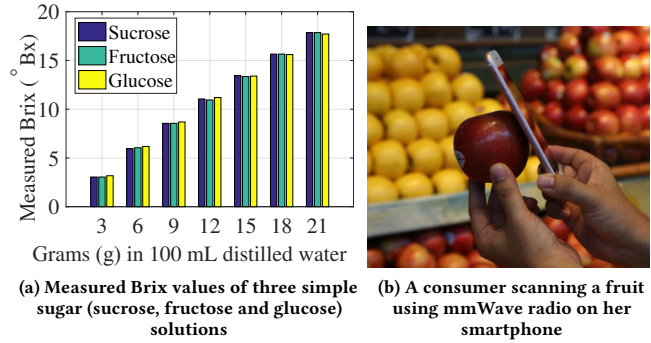**(b) A consumer scanning a fruit using mmWave radio on her smartphone**

Figure 1: (a) Brix can be used to measure the simple sugars (sucrose, fructose and glucose) commonly found in fruits; (b) Our propose mmWave reflection technique can be used by a consumer to determine SSC in fruits using smartphone

theoretical reasoning supported by experimentations and Section 4 studies the impact of different factors like size, shape and density on reflections. We develop our feature space in Section 5 and evaluate our system with experiments in Section 6. Section 7 discusses different limitations and potential of the proposed technique. We conclude in Section 9.

2 ESTIMATING SSC USING MMWAVE: FEASIBILITY STUDY

In this section, we show that it is feasible to use mmWave reflections to estimate SSC. We first present a metric that is commonly used to measure SSC and then demonstrate how mmWave reflections can be used to predict the metric.

2.1 Estimating SSC using Brix

Brix ($°Bx$) is a commonly used metric to measure soluble solid content in an aqueous solution. It represents the strength of the solution as the percentage by mass [54] where 1 $°Bx$ refers to 1 gram of sucrose in 100 grams of aqueous solution. Brix is widely used in wine, honey, and carbonated drinks industries.

A challenge in directly using Brix for measuring SSC in fruits is that the simple carbohydrates in fruits include not only sucrose but also fructose and glucose [42]. We first investigate if the sugar content of fructose and glucose solutions follow the same Brix scale which is originally proposed for the sucrose solution. We use 100 mL distilled water and the same portion of fructose, glucose, and sucrose masses for the Brix of 3 ∼ 21g with a step of 3g at room temperature of 25°C (following the sucrose to Brix conversion from [6] and temperature adjustment [64]). As we elaborate later, we use a plastic container instead of a glass container because a glass layer can attenuate 60 GHz signal by more than 5 dB [25]. An ATAGO PAL-BX/ACID F5 device [4] shown in Fig. 2 is used to measure the Brix values of these three simple sugar solutions. The resolution of this device is 0.1°Bx and the accuracy is ±0.2°Bx. These specifications are common in many commercially available digital portable refractometers [14, 31]. Fig. 1a shows that three simple sugars share the same scale of Brix measurement, which means that Brix can be used to measure the total SSC including sucrose, fructose and glucose. This makes Brix a suitable choice in our work where we are interested in measuring SSC in fruits.

It is also essential to understand how the Brix metric is related to user-perceived sugar content. Past researches have shown a
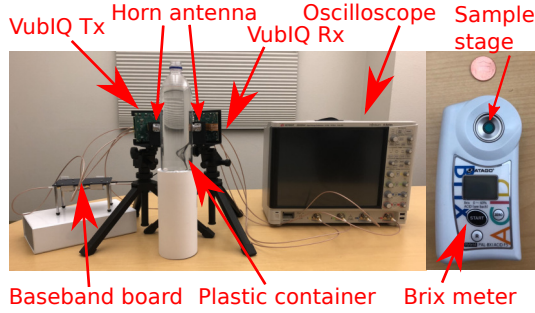
Figure 2: Our 60 GHz mmWave testbed using horn antenna and Brix meter

Figure 3: Block diagram of our 60 GHz testbed

Figure 4: Relationship between Brix and RSS for three simple sugar solutions

strong correlation between the Brix value and user-perceived sugar content. The perceived SSC is measured using the degree of liking metric which is measured through a hedonic score system (from 1 to 9 where 1 is dislike extremely and 9 is like extremely).

2.2 Estimating Brix using mmWave Reflections

Given that Brix metric can quantify SSC, we are now interested in investigating its relationship with 60 GHz mmWave signal reflections. This relationship can be used by a consumer to estimate the Brix of a fruit. Fig. 1b shows an illustration where a consumer uses his smartphone to transmit 60 GHz mmWave signal probe to a fruit, receive a reflection, process the reflected signal and estimate the Brix. It is worth noting that in this work, we assume that the 60 GHz signal probe is physically in contact with the fruit (as shown in Fig. 1b), and there is no signal path-loss between the two.

Although our intended use case is a mono-static (transmit and receive on the same device, i.e. smartphone) probing system, currently there is no commercially available 60 GHz transceiver systems in public that can be directly used for reflection analysis. Devices such as Netgear nighthawk x10 AD7200 and Tp-link Talon AD7200 WiFi access points are equipped with 60 GHz chipset and phased antenna array. However, the reconfigurable or programmable options of these devices are not available to the public, and the measurements for RSS and raw wireless signals are not allowed. Recently, several 60 GHz testbeds [39, 53, 56, 57] were built for developing better networking and sensing techniques. Since our objective in this work is to understand the feasibility of Brix estimation using 60 GHz reflections, we develop our own 60 GHz mmWave software radio testbed which acts as a bi-static (separate transmitter and receiver devices) probing system. Fig. 2 shows our experiment setup and the equipment. Fig. 3 shows the diagram of our 60 GHz testbed. At the transmitter side, we use VubIQ PEM-004 board [51] to generate the baseband sine wave signals and use VubIQ PEM-009 60 GHz RF front-end development kit [51] to up-convert the baseband signal to 60 GHz signals. At the receiver side, a Keysight DSOS254A oscilloscope is fed with four differential IQ signals. Two horn antennas with 20° 3-dB beamwidth are attached to the waveguide [51]. Note that although our presented analysis is based on the software radio testbed, it can be integrated to commercial 60 GHz devices in the future as long as they allow measuring RSS and other related signal quality metrics. The signal measurements (RSS, amplitude, etc.) provided by the oscilloscope are streamed to a host computer.

We use our 60 GHz mmWave system and perform controlled experiments to infer the relationship between SSC and mmWave signal reflect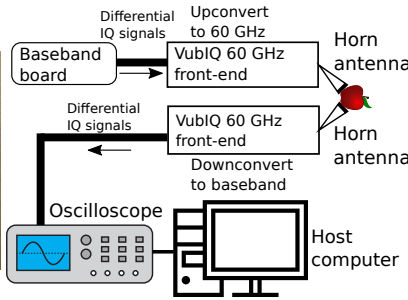ion. In the controlled experiments, we measure the signal reflection on three sugar (sucrose, fructose, and glucose) solutions as shown in Fig. 2. The reason why we use sugar solutions in place of actual fruits is that it allows us to control the sugar amount (thus the Brix). Since the fruits have very high water content with variable amount of SSC, the sugar solutions well approximate the reflection behavior in controlled settings. Of course, fruits have different density, surface scattering, penetration property, etc. which we will study in Section 4.

Fig. 4 shows the relationship between 60 GHz reflection Received Signal Strength (RSS) and the Brix values of the solutions. For each of the 8 Brix samples, the experiments are repeated five times at 25°C room temperature to achieve 90% confidence intervals. We observe that RSS decreases as the Brix value increases. The linear correlation coefficients of sucrose, fructose, and glucose are calculated to be 0.89, 0.96 and 0.86, respectively. This observation validates that 60 GHz signal reflection is influenced by the amount of soluble sugar content and it is feasible to be used as a sensing modality for inspecting SSC in fruits. We note that Fig. 4 shows a correlation between mean RSS and Brix, however, mean RSS alone is not sufficient to accurately estimate Brix (overlapping error bars). We show in Sections 5 and 6 that when the mean RSS is augmented with other statistical features calculated over RSS time series data and raw reflected signal, the Brix estimation accuracy substantially increases. In the next section, we explore the underlying reasons behind why the RSS changes with varying levels of SSC.

3 PERMITTIVITY AND MMWAVE REFLECTION

Varying levels of SSC affect the permittivity of the solution which in turn changes how much signal is reflected. In this section, we explain the impact of permittivity on signal reflection and elaborate on the relationships among permittivity, density, and Brix. As shown in Fig. 5, our Tx and Rx are in contact with the fruit while measuring RSS. Hence, we assume the signal propagation based on the near-field effects.

3.1 Reflection Coefficient

An object's permittivity indicates how much signal penetrates the object and how much reflects from it. Aside from the penetration and reflection, the signal is also absorbed and scattered by the objects. We let $L_r(\epsilon)$ denote reflection loss from the object with *relative permittivity* of ϵ. Relative permittivity is a ratio of the absolute permittivity of the material to the vacuum permittivity.

The reflection loss $L_r(\epsilon)$ can be calculated using the *reflection coefficient* (R). The reflection coefficient is the ratio of the amplitude
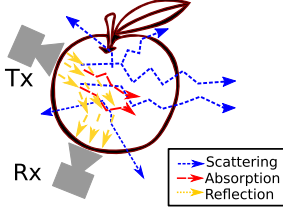
Figure 5: mmWave signal transmission, scattering, reflection, and absorption in fruits
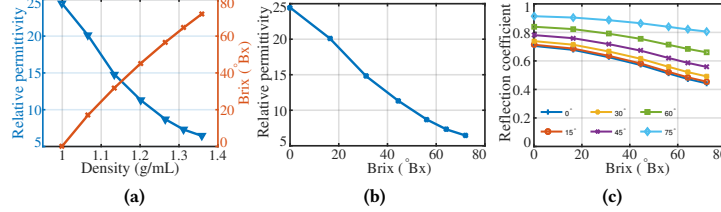

Figure 6: (a-b) Relationship between density, Brix and permittivity (c) The relationship between Brix and reflection coefficient remains unchanged for different incident angles
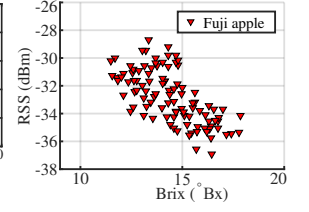

Figure 7: RSS variation for Fuji apple samples with different Brix values

of the reflected electromagnetic wave to that of the incident wave, and it can be calculated as [46]

$$R = \frac{1 - e^{-j2\delta}}{1 - R_i^2\, e^{-j2\delta}}\, R_i, \qquad \text{for } i \in \{\perp, \parallel\} \tag{1}$$

where $\delta = \frac{2\pi l}{\lambda}\sqrt{\epsilon - \sin^2\phi}$, l denotes the thickness of the reflecting source, λ denotes the signal wavelength and ϕ is the incident angle. $R_\perp$ and $R_\parallel$ are the Fresnel's reflection coefficients when the electric field is perpendicular and parallel to the incidence plane, respectively [46]. The reflection coefficient (R) can be used to estimate the reflection loss $L_r = P_O/P_I = |R|^2$ where P_O and P_I are the values of reflected (after reflection) and incident (before reflection) power.

3.2 Permittivity, Density and Brix

When SSC concentration increases in a solution/fruit, its permittivity and Brix value change. The change in permittivity results in how much signal is reflected (different reflection coefficient). Our proposed system attempts to infer Brix using the reflection (dashed arrow). We next explore the relationship between SSC density and permittivity, and SSC density and Brix. As mentioned before, since sugar water solution closely matches the characteristics of fruits, we use it as the basis of our explanation below.

3.2.1 SSC Density and Permittivity. Given the SSC density (the amount of SSC, i.e. sugar) of a solution, permittivity can be calculated using three existing models developed in previous researches. The models are developed primarily for fructose and are described below:

(1) Debye model: The model presents the permittivity of pure water [13] as

$$\epsilon_D(\omega) = \epsilon_\infty + \frac{\epsilon_0 - \epsilon_\infty}{1 + i\omega\tau_D} \tag{2}$$

where ϵ_0 is the static permittivity, ϵ_∞ is the infinite frequency permittivity, τ_D is a constant relaxation time, and ω is the angular frequency.

(2) Cole-Cole model: The model represents the permittivity of fructose solution when fructose amount is 1 monomol/L [10]:

$$\epsilon_C(\omega) = \epsilon_\infty + \frac{\epsilon_0 - \epsilon_\infty}{1 + (i\omega\tau_C)^{1-\alpha}} \tag{3}$$

where α is a parameter to define spectral shapes.

(3) Combined Debye and Cole-Cole model: The combined model is the Cole-Cole model function with the addition of a Debye term presenting the permittivity of fructose solution when fructose amount is more than 2 monomol/L [17]:

$$\epsilon_{CD}(\omega) = \epsilon_\infty + \frac{\Delta\epsilon_C}{1 + (i\omega\tau_C)^{1-\alpha}} + \frac{\Delta\epsilon_D}{1 + i\omega\tau_D} \tag{4}$$

where $\epsilon_0 = \epsilon_\infty + \Delta\epsilon_C + \Delta\epsilon_D$.

Based on these three models, we calculate the permittivity for a given solution with varying SSC densities. The relationship is shown in Fig. 6a. We observe that as the SSC density increases, the permittivity decreases monotonically. The density values shown here correspond to typical Brix values observed in fruits. The decrease in permittivity forms the basis of our proposed technique.

3.2.2 SSC Density and Brix. Calculating the Brix value given the SSC density is also challenging due to the dependence of Brix on solution temperature. Even for a given fructose solution with a fixed density, Brix value can vary at different temperatures. There is not ready-to-use lookup or conversion table available for different room temperatures other than 20°C. We address this issue by allowing a Brix adjustment process as described below.

The Brix value at 20°C can be calculated using the inverse function version of Kimball equation in [28] as

$$B_{20} = 412.838\sqrt{\ln\left(\frac{\rho_{20}}{0.524484}\right)} - 330.872 \tag{5}$$

where B_{20} is Brix value at 20°C, and ρ_{20} is the density of the solution at 20°C. ρ_{20} can be calculated as

$$\rho_{20} = \rho_{20}^{pw}\Gamma_{20} = \rho_{20}^{pw}(\Gamma_T - \alpha_{20}^T) = \rho_{20}^{pw}(\rho_T/\rho_T^{pw} - \alpha_{20}^T) \tag{6}$$

where ρ_{20}^{pw} refers to the density of pure water at 20°C, Γ_{20} is Specific Gravity (SG) of the solution at 20°C, Γ_T is the SG of the solution at temperature T, α_{20}^T is a constant adjustment of SG values between the temperature T and 20°C, which is formalized from Table 18 in [44], ρ_T refers to the density of solution at the temperature T and ρ_T^{pw} denotes the density of pure water at temperature T. We also formalize constant adjustment of Brix values between temperature T and 20°C using β_{20}^T as

$$B_T = B_{20} + \beta_{20}^T. \tag{7}$$

Therefore, the Brix value at the temperature T, B_T, based on the density of solution ρ_T can be derived based on Eq. 5-7 as

$$B_T = 412.838\sqrt{\ln\left(\frac{f(\rho_T)}{0.524484}\right)} - 330.872 + \beta_{20}^T \tag{8}$$

where $f(\rho_T) = \rho_{20} = \rho_{20}^{pw}(\rho_T/\rho_T^{pw} - \alpha_{20}^T)$. Because ρ_{20}^{pw}, ρ_T^{pw}, and α_{20}^T can be regarded as constants [6, 52], we rewrite Eq. 6 as a function of ρ_T in Eq. 8. Fig. 6a shows the relationship between the SSC density and the Brix value calculated using Eq. 8. Fig. 6b also shows that permittivity monotonically decreases as the Brix value increases. This relationship in essence allows our system to estimate Brix using the reflection.

Fig. 6c shows how reflection coefficient changes with Brix for different incident angles (ϕ in Eq. 1). We find that the relationship (monotony property) does not change over different incident angles, making our proposed system more robust to such changes.

Container	Diameter (mm)	R_a in mm
Type A	69.43	Smooth (0.0028)
Type B	89.01	Smooth (0.0030)
Type C	105.43	Smooth (0.0027)
Type D	75.98	Coarse (0.088)

Table 1: Four types of containers used in controlled experiments, and their size and shape properties

4 ESTIMATING SSC IN FRUITS

We now turn our focus on estimating SSC in real fruits and understanding different factors that affect the estimation accuracy. We use our mmWave probing system on 100 different Fuji apples and collect the reflection RSS values. The RSS values along with the Brix values are shown in Fig. 7. Compared to Fig. 4, the generic trend of RSS decreasing with higher Brix is also observed with the apple samples. This means that the mmWave reflection characteristics that we discussed previously for sugar solutions are maintained for the real fruits. However, we observe that the linear correlation coefficient is ≈ 0.65. In many cases, apples which have the same Brix value have considerably different RSS. This means that compared to the sugar solution, estimating Brix using mmWave reflections on fruits need to involve a range of other factors.

When mmWave signal impinges a fruit, it is not considered as a uniform medium because the signal reflection and propagation inside it is complex due to different fruit characteristics. The signal undergoes scattering, reflection, and absorption as shown in Fig. 5. We assume they are based on the near-field effects. Such signal propagation is dependent on many different factors: (i) fruit diameter, (ii) fruit surface roughness, (iii) flesh density and (iv) flesh structure. Diameter and surface roughness depend on the size and shape of a fruit, while the flesh density and structure are unique characteristics of different types of fruits.

4.1 Impact of Fruit Size and Shape Irregularities

Compared to previous experiments (sugar solution in a plastic container), fruits have highly variable, imperfect shapes and sizes. Even the fruits of the same cultivar (e.g., Fuji apples), these properties can vary significantly. Since these factors are difficult to accurately measure, we use controlled experiments with sugar solution in plastic containers to understand their impacts. We use two different types of solutions (10 grams, and 20 grams of sucrose in 100 mL water) to evaluate the mmWave reflections. We choose four different types of plastic containers, and their diameter and surface roughness values (measured using caliper) are shown in Table 1.

The thin plastic wall of the containers does not imitate fruit skin. In this work where the fruit skin is not considered a separate layer compared to the flesh of the fruit. Other than the fact that studying the reflection properties of the skin is challenging with our testbed setup, our objective is to design a non-destructive method of Brix estimation where the user does not have to alter the fruit (e.g., remove skin) to estimate its Brix. Typically, plastic containers are made of polyethylene terephthalate (PET) [3]. The four plastic containers we use in our study have wall thickness within the $0.34-0.38mm$ range. Based on the permittivity of PET at 60 GHz [11] as well as the Equations 1, its reflection coefficient can be calculated as 0.067 for a wall thickness of 0.34 mm and 0.080 for a wall thickness of 0.38 mm under small incident angles ($0°$ to $5°$). As a reference,
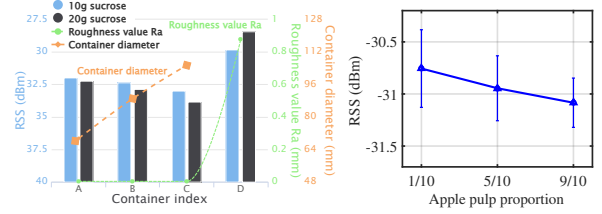

Figure 8: Impact of diameter and roughness on RSS

Figure 9: Relationship of different pulp portion and RSS (Brix set to 13.2)

the reflection coefficient of air is 0 at 60 GHz [43]. Such a low value of reflection coefficient suggests that signal impinging on the plastic container with these properties mostly penetrates through the wall. Since sucrose molecules interact less with water molecules, creating non-uniform reflections and complex permittivity, sucrose solution is considered not uniform at molecule levels [32, 47]. This means that most of the reflections occur from the solution inside the container, making such a setup suitable for our study. Given that container thickness does not have a significant impact on reflection, we use the containers to study the impact of diameter and roughness. We repeat the experiments for 12 containers (3 containers for each of the four types) to calculate mean values.

- **Diameter:** Fig. 8 shows how RSS varies with the change in container diameters. Containers A, B, and C have similar surface roughness, hence, variations in RSS for the same sucrose concentration is due to variation in their diameters. We find that the RSS decreases with the increase in diameter. Note that this result is difficult to generalize but our observations are useful in understanding the range of diameters typical in fruits (a few centimeters).

- **Roughness:** Fig. 8 shows a comparison between RSS for Containers D and the other three. We measure the roughness using the procedure described in [8]. Here, we measure the wall thickness of each container using a cantilever 20 times, and calculate the mean wall thickness ($\bar{Y}$). We then perform additional $N = 20$ measurements of thickness (Y_i) at randomly chosen locations on the container wall. The roughness is then calculated as $R_a = \sum_{i=1}^{N} |y_i|/N$ where $y_i = (Y_i - \bar{Y})/2$. These roughness values are shown in Table 1. Since the increase in diameter decreases the RSS, if Container D had the surface roughness similar to that of Containers A, B, and C, the RSS of the Container D would have been between that of Containers A and B (based on its diameter). However, Container D has the highest RSS which can be attributed to its coarse surface. The roughness of the surface is known to scatter the signal more [7], resulting in more reflection and higher RSS.

4.2 Impact of Flesh Density and Structure

Apart from the size and shape related factors discussed above, different fruits have different density depending on the percentage of water and non-soluble contents (e.g., fiber tissues). The non-soluble contents also significantly affect the absorption and scattering behavior of the mmWave signal. Authors in [33] showed that in case of mmWave signals, when the scatterer's particles are smaller than the wavelength of the mmWave signal, the absorption and scattering can be approximated using the Rayleigh model. However, when

Metrics	All features	GFS	GS	FS
Corr. coeff.	0.8177	0.8302	0.8294	**0.8397**
MAE	0.7723	0.7174	0.6885	**0.6812**
RMSE	0.9660	0.9349	0.9602	**0.8915**
NRMSE	0.0666	0.0644	0.0662	**0.0618**

Table 2: Correlation coefficient and prediction errors for three searching methods using the feature subset evaluator CFS
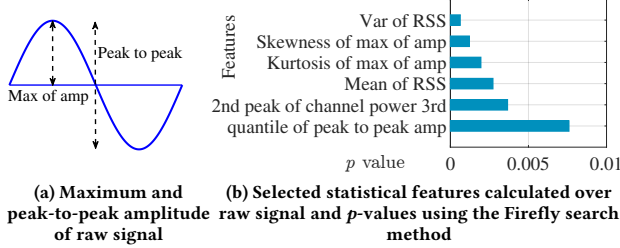

(a) Maximum and peak-to-peak amplitude of raw signal

(b) Selected statistical features calculated over raw signal and *p*-values using the Firefly search method

Figure 10: Statistics calculated over the raw reflected signal shows strong correlation with the Brix value
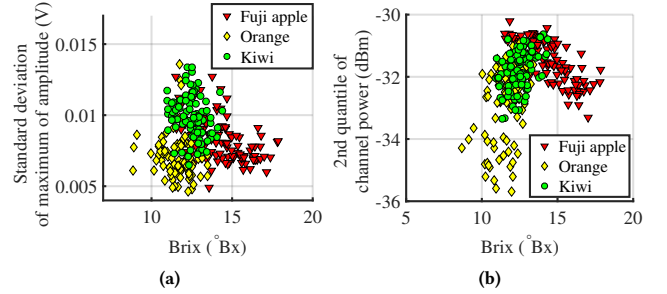

(a) (b)

Figure 11: Although the signal features are correlated to Brix, the differences in correlation for different fruit types motivates us to develop separate and common regression models for estimation

the size of the scatterer's particles is larger than 1 or 2 mm (as in fruit flesh structure), Rayleigh model based approximation does not always stand valid. In such a case, Mie Theory [36] is used to model the scattering behavior. In terms of absorption, Lambert-Beer Law [59] models mmWave absorption similar to the optical frequencies.

Since the density and structure of the fruit flesh is difficult to measure, we rely on the controlled experiments to understand the impact of flesh density on reflection RSS. We first produce pure pulp (no added water) of 20 apple samples and measure the Brix value of the pure pulp. We then prepare a sucrose solution which has the same Brix value as the pulp. Finally, the pulp and the sucrose solution are mixed in different proportions (1:9, 5:5, 9:1). Since the pulp represents the non-soluble content of a fruit, varying levels of it in sucrose solution allows us to imitate different flesh densities. This method allows us to study the impact of varying levels of flesh density while keeping the Brix value constant. Fig. 9 shows the RSS for the pulp solutions. It can be observed that as the pulp portion increases, the RSS reduces. This can be attributed to higher absorption and scattering resulting from more non-soluble content.

We find that various factors such as diameter, surface roughness, flesh density and structure affect the mmWave reflection. Fruits exhibit large diversity in terms of these factors. For example, in our experiments, Fuji apples and oranges have the mean density of 0.80g/mL while kiwis' mean density is 1.01g/mL. Similarly, apples, oranges, and kiwis have different diameters (75.12mm, 84.82mm and 53.85mm respectively). This means that it might be difficult to use one regression model to estimate their Brix using RSS. Hence, we propose to train and use a dedicated regression model for different fruits. We will compare the per-fruit type models and common models in our evaluation (Section 6). In the next section, we first describe our feature space and demonstrate how to develop the estimation models.

5 FEATURES AND BRIX ESTIMATION MODELS

We are now interested in developing a Brix estimation model for each type of fruit. To do so, we first explore statistical features of mmWave signal reflection.

As shown in Fig. 7, purely using RSS can yield lower Brix estimation accuracy. This is expected given that RSS is a coarse-grained

statistic calculated over the reflected signal. A time-series data can be further characterized by time and frequency domain analysis to understand a wide range of properties otherwise not available through mean. The usage of statistical features are widely found in literature for feature extractions from various types of time-series signals [27, 30]. We explore the use of fine-grained statistical features calculated over the raw reflected signal. Specifically, we consider three additional properties of the raw signal: (i) maximum value of amplitude, (ii) amplitude peak-to-peak value in time, and (iii) channel power in frequency domain. The first two features are calculated over time series (shown in Fig. 10a). Using these three properties, we calculate first, second and higher order statistics including mean, median, variance, standard deviation, skewness, kurtosis, etc. Note that these statistics are calculated individually for each of the three properties mentioned above and also for RSS to create a feature vector.

In order to evaluate if these features are useful in estimating Brix, we apply the correlation-based feature selection (CFS) [22] with the following three feature searching methods:

(1) Greedy forward search (GFS): The greedy process iterates through the feature space and at each step includes an additional feature that when combined with the selected subset provides the highest predictive ability.
(2) Genetic search (GS), which uses the genetic algorithm [20] to perform the searching procedure of feature selection.
(3) Firefly search (FS), which imitate the biochemical communication system of real fireflies using luminescent flashes as signals [55].

The output of these searching methods is a subset of features, which are highly correlated to Brix but less correlated with each other. The feature dimension is consequently reduced. The procedure is applied on 100 Fuji apple samples (same as in Section 4). We use three metrics - Mean average error (MAE), Root mean squared error (RMSE), and Normalized root mean squared error (NRMSE) - to evaluate the feature selection performance of these methods by assessing their prediction errors. The metrics can be calculated as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i - p_i)^2}; \quad \text{NRMSE} = \frac{\text{RMSE}}{\frac{1}{n}\sum_{i=1}^{n} m_i} \quad (9)$$

where n is the number of samples, m_i and p_i denote the measured value and the predicted value of the i-th sample, respectively. Since RMSE performs the quadratic operation before the arithmetical average operation, large errors are amplified. Therefore, RMSE is able to capture large errors with a relatively higher weight compared

Figure 12: Our 60 GHz mmWave testbed setup showing how we measure signal reflection from a fruit

to MAE. NRMSE is able to avoid the scale dependency of errors in RMSE, but keep the property of sensitivity to large errors like RMSE. The prediction is considered excellent when NRMSE is less than 10% [26].

Table 2 presents the results of linear regression using the three feature selection methods compared to using all features. We observe that with feature selection, there is a minor improvement in the correlation coefficient and a minor reduction in prediction errors. Due to its high correlation coefficient and lowest prediction errors, we choose Firefly algorithm as our feature search method.

Different searching algorithms have different searching strategies resulting in a subset of features that are different from each other. During the feature selection procedure, the dimension of feature space is reduced from 76 to 33, 48, and 44, using the three searching methods, greedy forward search, genetic search, and Firefly search, respectively. We observe that the larger number of features in the subsets using genetic search or Firefly search keep more features in frequency domain than greedy search. In Section 6, we develop regression models that use two sets of features: (i) the chosen subset based on feature selection and (ii) RSS-based features only. This is because many of the commodity wireless radios usually record RSS which can be easily extracted to user-space for estimation.

We also investigate the subset of features using *p*-value assessment. *p*-value is commonly used to present the significance of variables/features in multiple regression. We propose a hypothesis that each of the features does not help in predicting the Brix value. Smaller *p*-value of a feature means higher probability to reject the null hypothesis, that is, the feature is important to predict the Brix. Typically, the significance level is 0.05 for *p*-value [37], which mean that if *p*-value is above 0.05 then the null hypothesis is not rejected. Fig. 10b shows that some features in the feature subset have very low *p*-values. Let us take the variance of RSS (first feature in Fig. 10b) for an example. As Fig. 5 shows, the signals undergo scattering and reflection behaviors inside the fruit sample, and the received signals will be a mixture of the signals from various propagation sources (scattering, diffusion, reflection). As a result, the signal fluctuation will be non-trivial [1]. Variance of RSS shows the variation of received signals, and thus is able to represent this signal fluctuation behavior.

We next evaluate if these features can also be used for Brix estimation in other fruits. For this purpose, we collect mmWave reflection for 100 oranges and 100 kiwis along with their Brix values in addition to the 100 Fuji apple samples. Fig. 11a shows standard deviation of maximum of amplitude and its relationship with Brix for all 300 different fruit samples. It can be observed that this feature is noticeably different for different type of fruits. This can be attributed to the heterogeneity in fruit diameter, fruit surface roughness, flesh density and structure of different fruits along with
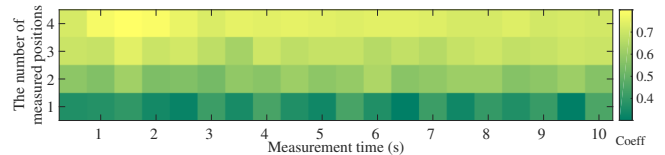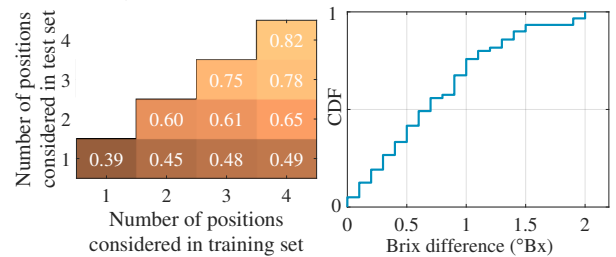

Figure 13: Measuring a fruit sample at multiple positions results in a more accurate estimation of Brix, however, measuring for a longer duration does not improve the accuracy

(a) Correlation coeff. with varying number of measured positions considered in training and testing

(b) CDF of variations observed in measured Brix values between the pieces of the same fruit sample

Figure 14: The number of measured positions/orientations and its impact on estimated Brix

their Brix value. Another feature, second quartile of channel power, is depicted in Fig. 11b. We find the feature provides clear distinguishability in terms of its relationship with Brix. However, the relationship between the feature and Brix is different for different fruit types, further making a case that one Brix estimation model common across all fruit type can also be developed albeit at some penalty in accuracy.

6 EVALUATION

Since there is no commercially available off-the-shelf mmWave transceiver device that can provide in-band sensing control, we use our testbed described in Section 2 for our evaluation. Fig. 12 shows the experiment setup with horn antenna close to a fruit sample. The incident angle is set to 45°. We proceed with an exhaustive evaluation using 100 apples, 100 oranges, and 100 kiwis. We select these three types of fruits because of their diverse shapes, sizes, surface roughnesses (e.g., kiwi vs. apple), flesh structures (orange vs. apple) and densities. Additionally, these fruit types provide significant diversity in terms of their Brix values, allowing us to better study our estimation accuracy and sensitivity (as discussed later in Fig. 16a). Each fruit sample is measured from 4 different positions (every 90°) to obtain multiple observations. At each position, mmWave signal reflections are measured for 10 seconds. After mmWave measurement, ATAGO PAL-BX/ACID F5 device [4] (shown in Fig. 2) is used to measure the Brix value of the fruit as the metric of its sugar content. We cut every fruit sample into small pieces and use a fruit press to acquire its pure juice. The juice is then put into the sample stage of the Brix meter (in Fig. 2). After every measurement, we clean the sample stage using distilled water to remove any remaining juice or residue, and dry it using lab paper to assure the next measurement is not affected by the previous measurement.

We apply three different regression algorithms to evaluate the estimation performance: (1) linear regression (LR), (2) random forest regression (RFR), and (3) support vector regression (SVR) with radial basis function (RBF) kernel. *Random Forest* is a popular machine learning algorithm for classification and regression [9]. It constructs multiple decision trees to perform the majority of votes

(a) Results of correlation coefficients using different regression algorithms

(b) MAE of different regression algorithms

(c) RMSE of different regression algorithms

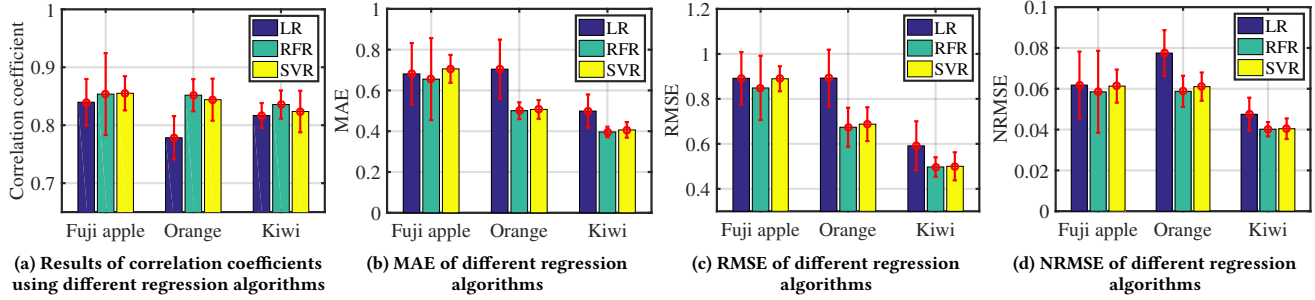(d) NRMSE of different regression algorithms

Figure 15: (a-d) Regression model developed using the subset of features (RSS + Amplitude) provide high Brix estimation accuracy (higher correlation coefficient and lower MAE, RMSE, and NRMSE)

for classification, or to calculate a mean for regression. *Support Vector Machine* is another popular machine learning algorithm designed for classification. Nevertheless, it can also be applied for regression [16]. Aside from keeping the advantages of kernel function, the major difference in the usage of classification or regression is that the margin created by hyperplanes contributes a "tube" area, where the cost function excludes the training points that are beyond the margin.

6.1 Impact of Measurement Time and Positions (Usability)

Before evaluating the relationship between Brix and the reflection profile of the fruit samples, we first evaluate how many positions one needs to measure and for how long in order to yield an accurate estimate of Brix. This reflects how usable our proposed scheme is for a user and how much effort is necessary for evaluating one fruit from a consumer's perspective. The dataset of Fuji apples is used in this evaluation. We select the linear regression model and do 5-fold cross-validation on the training set of 100 Fuji samples. Fig. 13 shows the correlation coefficient for the linear regression and its relationship with the number of observed positions and time. Here, we train the regression model using the same number of measured position as it is used in testing. For example, when the correlation coefficient for 2 measured positions is evaluated, the number of measured positions used in training is also 2. As we can see, as the number of observed positions increases, the coefficient also increases. On the other hand, we observe that measuring for a longer time does not necessarily improve the estimation accuracy. We find that in most cases, it is sufficient for a user to measure a fruit for approximately 1.5 seconds at 4 different positions to yield a correlation coefficient of around 0.82.

Fig. 14a shows the correlation coefficient of Brix estimation when the different number of positions are used for training and testing. Here, we only focus on cases where the number of measured positions used in training are greater than or equal (diagonal - same as Fig. 13) to the number of measured positions in testing. The objective here is to understand whether training using more number of measured positions improves the Brix estimation or not. We observe considerable improvement in Brix estimation accuracy for all combinations. For example, in the case where 4 measured positions are used for training compared to 1 measured position, the correlation coefficient increases from 0.39 to 0.49 when 1 measured position is used in testing. This demonstrates that the impact of variations observed in measurements from different orientations can

be reduced by training with model with more number of measured positions.

The variations observed in estimated Brix can also be attributed to how the ground truth of Brix is measured. In a typical fruit sample, Brix value varies significantly within the same fruit. To evaluate the inherent Brix variation within the fruits, we take 20 Fuji apple samples and cut each of them in four equal-sized pieces. We then separately measure the Brix value of each fruit piece. For each piece, we measure the Brix value of each piece three times. We observe that the Brix value of the same fruit piece (three measurements) varies from $0.0 \sim 0.2°$Bx (same as the accuracy range of our Brix meter). We calculate the mean of the three measurements and calculate the differences of means for the four pieces of each fruit sample. For a total of 20 Fuji samples, this results in 120 (20×6) Brix difference values. Fig. 14b shows how much the measured Brix value varies between two pieces of the same fruit sample. This presents that even within the same fruit, the measured Brix values vary significantly (standard deviation of $0.5232°$Bx and the maximum observed variation of $2°$Bx.

In the results presented in the paper, we only measure one Brix value of the entire fruit which can be treated as the mean of the Brix value of different parts of the fruit. When the mmWave reflections are measured from different orientations, the Brix value of that portion of the fruit can be different from the mean value, leading to a lower correlation coefficient. Hence, when multiple measurements are performed, the estimated Brix value appears closer to the measured ground truth mean Brix value.

6.2 Comparing Machine Learning Algorithms (Accuracy)

We use the 4 positions and 1.5 seconds of measurement time per position for each of the 300 fruit samples. Here, the LR is compared with RFR and SVR in terms of their estimation accuracy. We build separate estimation model for each fruit type and use 5-fold cross-validation for training and testing. Figs. 15a, 15b, 15c, and 15d show the correlation coefficient, MAE, RMSE, and NRMSE, respectively, for all three fruit types and three regression models. Users can taste a difference between apples when more than $1°$Bx difference occurs [24]. Since our prediction MAE and RMSE are within this range, our accuracy boundaries are able to meet user's expectations. The highest mean correlation coefficients (apples: 0.86, oranges: 0.85, kiwis: 0.84) are observed with RFR, while SVR-based model performs similar to the RFR model. Using RFR and SVR, the model better adapts to the available Brix values in the training set which results in better estimation accuracy using these algorithms. It can

(a) Samples of three fruit types have different range of Brix

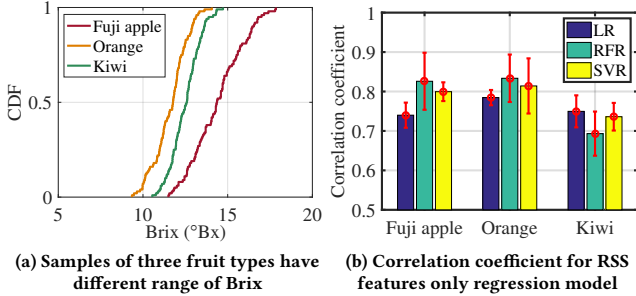(b) Correlation coefficient for RSS features only regression model

Figure 16: (a) Different range of Brix for different fruit types affect the accuracy performance of different regression algorithms, (b) Regression models developed only using RSS features achieve reasonably accurate estimation

Features	Metrics	Cross validation					
		5-fold			10-fold		
		LR	RFR	SVR	LR	RFR	SVR
All features	Corr. coeff.	0.67	0.71	0.68	0.68	0.72	0.69
	MAE	0.96	0.79	0.90	0.97	0.83	0.89
	RMSE	1.18	1.07	1.21	1.17	1.10	1.23
	NRMSE	0.092	0.083	0.094	0.091	0.086	0.096
RSS features only	Corr. coeff.	0.65	0.68	0.65	0.65	0.66	0.64
	MAE	0.91	0.93	0.91	0.91	0.92	0.92
	RMSE	1.22	1.17	1.20	1.22	1.20	1.26
	NRMSE	0.094	0.091	0.093	0.095	0.093	0.098

Table 3: Correlation coefficient, MAE, RMSE, and NRMSE for estimation models developed using samples of all fruit types

be observed that LR performs poorly in case of oranges and kiwis. This can be attributed to the fact that our apple samples have a larger Brix range as shown in Fig. 16a compared to that of oranges and kiwis. This observation is in line with Fig. 11 as well. This larger Brix range also causes the relatively higher MAE and RMSE values compared to those of oranges and kiwis. It is worth noting that even though RMSE of Fuji apples is higher, all the NRMSE values across the three fruit types are far less than 10% [26], further validating the accuracy of our proposed technique.

6.3 RSS Only Vs. All Features

As we noted before, most commodity radio chipsets usually output only RSS value to the user-space. Profiling raw signal requires low-level (driver and firmware) control in most wireless radio devices. Recent advances in making Channel State Information (CSI) [23] available in user-space has resulted in a plethora of sensing related research. However, this has been limited to 2.4/5 GHz WiFi devices. Such modification and development of tools are likely for 60 GHz mmWave WiFi in the future. Hence, we compare the accuracy of Brix estimation for models developed for just RSS-based features and all features. Fig. 16b shows the correlation coefficient of regression models developed only using RSS features. We find that just utilizing RSS can still achieve reasonably high estimation accuracy (for both RFR and SVR). Additional features based on amplitude can certainly provide a more fine-grained look at the variations in reflection and can in turn achieve higher accuracy.

6.4 Common and Cross Models

We now evaluate the Brix estimation accuracy for two different cases - common models and cross models. In common models, we combine the observations of all fruit samples (300 fruits × 4 positions per fruit) and train a type-agnostic model. We then apply 5 and 10-fold cross-validation, and evaluate how the model trained using observations from all different fruit performs in estimating Brix of given fruit instances. The results are presented in Table 3. The observed correlation coefficient is 0.67, 0.71 and 0.68 for LR, RFR and SVR, respectively. We find that due to the diversity of fruit characteristics and their relationship with reflected RSS, the common model performs poorly in SSC estimation. The MAE and RMSE values are worse here as expected. This further validates our claims in Section 4 where we studied the impact of these factor in controlled experiments.

We also evaluate cross models where a regression model is trained for one fruit type and used for testing on another. Fig. 17 presents the confusion matrix for each regression algorithm in case of the cross models. Similar to the common model, we find that LR performs better compared to RFR and SVR because the latter two suffer from over-fitting. Such over-fitting is not observed in per-fruit models. The common and cross models confirm that the diversity in fruit characteristics necessitates development of per-fruit type model for higher estimation accuracy. There is an opportunity to develop fruit-type-agnostic models which can be applied to test any fruit type. However, such a model requires large, diverse training dataset and through validation. We leave this exploration to future work.

6.5 Fruit Type Identification

Lastly, we evaluate how our mmWave signal probing system can be used to identify the type of fruit. Such identification can be useful for developing machine intelligence with applications in object recognition, smart grocery scales and automatic fruit sorting. We use the same feature space to develop classification models using Naive Bayes, Random Forest, and Logistic Regression Classifier. The accuracy of fruit-type identification with 1200 samples and 10-fold cross validation is found to be 84.67%, 92.33%, and 92.67% respectively. This suggests that our proposed approach has the potential to not only identify the Brix value of a fruit but also to identify the type of fruit based on mmWave reflections.

6.6 Impact of Temperature and Other Atmospheric Variables

All experiments presented till now were performed when the fruit sample and the sugar solutions are at room temperature (25°C). Temperature is known to have a non-trivial impact on the dielectric property of reflecting objects, altering the amount of radio frequency signal reflected back. To study this impact, We consider a sucrose solution with different sucrose densities, 15g/100mL, 18g/100mL, and 21g/100mL. We vary the temperature of the solutions to consider three levels: cold (12°C), room (25°C) and warm (37°C). Fig. 18 shows the RSS measurements of the three solutions at different temperatures. We observe that (i) the underlying trend of decrease in RSS with increases in sugar content remains consistent even with variations in temperature, (ii) the absolute RSS value at different temperatures for different sugar contents decreases as the temperature increases. This means that in order to use our proposed technique of SSC estimation for fruits at different temperatures,
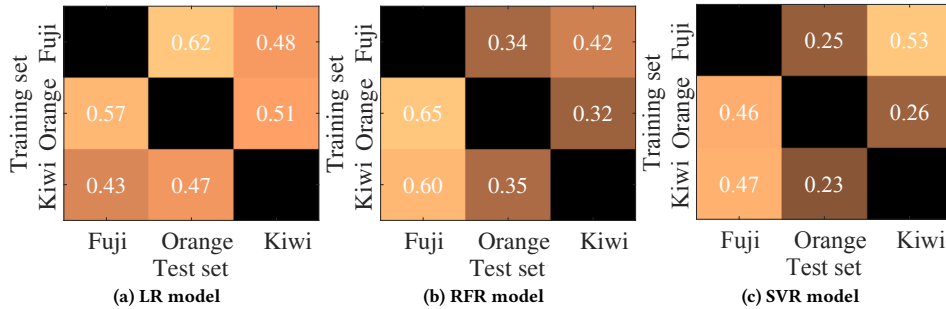
Figure 17: Correlation coefficient decreases significantly when applying regression model trained using samples of one fruit type to others

(a) LR model (b) RFR model (c) SVR model

Figure 18: RSS results of three sucrose solutions under three different temperatures

temperature itself should be included in the training of the model. Note that the Brix meter used in our study can perform automatic temperature compensation when operating at different temperatures, but it requires recalibration using pure water. Hence, the model should be trained using the RSS features as well as the recalibrated Brix values. Since the underlying relationship between RSS and SSC is maintained, such a model can still accurately estimate SSC at different temperatures.

It is worth noting that other factors such as humidity in the air, moisture in the fruit and dirt on the surface of the fruit can also affect the mmWave reflections. Since we assume that the 60 GHz mmWave probing device is in contact with the fruit when estimating SSC, the atmospheric humidity is likely to have a little impact on the reflections. Variations in amount of moisture in the fruit can certainly have a non-trivial impact. For the different types and cultivars of fruits, moisture can be measured and included in the training model. As our primary application is a consumer determining SSC while shopping, we assume that the fruits do not have dirt on their surface. However, if our proposed SSC estimation is to be applied to do the same in other settings (e.g., farms), the impact of dirt and other chemicals on the fruit surface should be studied. We leave the exploration of these factors and their impacts to future work.

7 LIMITATIONS, POTENTIAL AND FUTURE WORK

Limitations. As mentioned before, our proposed technique of estimating SSC in fruits is only the first step towards developing a new set of consumer-oriented food and agriculture produce sensing techniques. We show that it is indeed feasible to use mmWave wireless radio chipset available in consumer devices to sense the sugar content in fruits. However, we believe that there are multiple limitations of our work:

- Fruit types and cultivar: Even for three types of fruits that we investigated in our evaluation, a diverse set of fruit characteristics (size, shape, flesh structure, etc.) are observed. We have tried to characterize the impact of these characteristics on reflection using sugar solution studies. Also, our proposed approach does not take fruit cultivars (e.g., Gala vs. Fuji apples) in account.
- Need of larger dataset to combat heterogeneity: Fruit produce available to today's consumers largely depends on the supply-chain operations. It is difficult to assess the diversity of our fruit samples compared to the ones available at different part of the country or world. Extensive experimentation

is required to understand the diversity of fruits and how they impact our Brix estimation framework.
- End-user studies: Due to the controlled nature of our experiments, we are unable to understand different factors that can affect the accuracy and usability when our system is used by end users. For example, it is not clear what are the user's preferences in terms of measurement time on a fruit. We plan to study this factors in our ongoing work where a handheld mmWave probe can be used to evaluate its real-world performance.

Potential. In spite of the limitations, we believe that our proposed research opens a new direction for a variety of consumer-oriented and other sensing techniques. We describe a few of them below.

- Sensing sugar in other food and beverages: Our proposed approach can be extended to drinks and beverages like wine and dairy products. A consumer can also evaluate the amount of sugar in beverages where nutrition information is not available.
- Augmenting supply-chain operations: With development of low-cost Brix estimation device, each entity (not just the farmers) involved in bringing the produce from farms to consumers can assess the exact time at which the produce should reach the consumers for highest satisfaction. Such estimates can simplify operations and reduce storage cost. Even for a consumer, this information can provide an accurate estimate on "time to ripe", and consuming a produce at the desired level of SSC can encourage a healthy diet.
- Smart groceries: Evaluation shows that our mmWave system can also classify the fruit type. Automatic fruit type and cultivar identification can considerably reduce manual labor and eliminate the need to paper tags on produces. This can not only reduce the maintenance cost of grocery store, but it can also make way for the development of new tools for automatic produce recognition at the weighing scale.

8 RELATED WORK

8.1 60 GHz mmWave Sensing and Networking

mmWave signals have been used as an effective sensing modality for a variety of applications. Authors in [53] designed an object motion tracking system using 60 GHz signals. Its beam scanning mechanism can locate the object's initial position, and then track its trajectory using a signal phase information. Instead of using phase information, an RSS based imaging system was presented in [63]. It leverages RSS measurements only to create an image of an
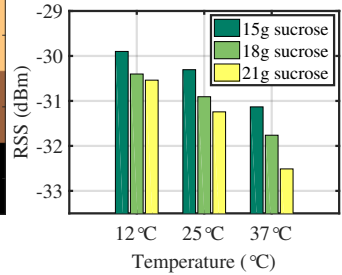
object, and precisely estimate the object's surface shape and contour. Another RSS-based object recognition system using a single 60 GHz mobile device was proposed in [62]. Beyond object recognition, 60 GHz is a promising technology for health monitoring. A vital sign monitoring system was designed in [57] to monitor breathing rates, heart rates, and sleep using the RSS of 60 GHz signals. Authors in [45] designed a blood glucose estimation system which exploits the signal penetration property of 60 GHz signals to measure the penetration on the thin skin between thumb and pointer finger.

60 GHz millimeter wave wireless communication is being used to design the next generation of WiFi networks. It is currently an active area of research. Authors in [49] provided link level characterization of blockage and reflection in an indoor environment. 60 GHz communication has also been studied for wireless links in data center networks [61], for wireless local area networks with the assistance of traditional 2.4/5 GHz WiFi [39], for mobile motion sensors [58], and for potential VR/AR applications [60]. With this ever-increasing interest in the technology, 60 GHz WiFi is likely to be ubiquitous in indoor WLANs in the coming years.

8.2 Fruit Quality Assessment Using Electromagnetic Waves

NIR spectroscopy is a popular non-invasive technology for fruit quality estimation [48] because it can measure important optical properties of fruits. During the past decade, many research approaches using this technology have been proposed to measure the constituents of various fruits. Authors in [38] presented estimation of the sugar content of Fuji apples using Fourier transform NIR spectroscopy, where the partial least squares (PLS) regression models between measured diffuse reflectance spectra and sugar content were evaluated by correlation coefficient at four difference sensing distances. The sugar content in oranges was studied in [35], where authors applied five spectral pre-processing methods at different spectral ranges to build regression models. Authors in [2] investigated the relationship between NIR spectroscopy measurements and the sugar content of kiwifruit at three different temperatures using discriminate analysis and PLS regression. The sugar content and firmness of two cherry cultivars were proposed in [34]. Although the performance of these techniques are found to be accurate, they require professional-grade, high-cost hardware setup. Due to the requirement of dedicated, custom hardware, these techniques are not readily accessible to consumers.

Some studies have proposed the use of a low-cost imaging imaging spectrometer [5]. For example, a low-cost hyper-spectral imaging system operating in visible and NIR bands is presented in [19]. One of its applications is to distinguish the ripeness of 10 various fruit types without the need of system calibration for each specific fruit. All of these approaches still require some additional device embedded into or attached to a smartphone, making them less suitable for end users. Instead, we demonstrate the feasibility of using 60 GHz transceiver embedded in smartphones to perform SSC estimation without any external attachment. As stated before, this enables a "reuse" of existing 60 GHz wireless networking hardware for a sensing application. Channel State Information (CSI) has been the research focus for many WiFi-related sensing papers. Our current platform can be extended in the future to get CSI. Given that RSSI might be more easily available on COTS 60 GHz devices,

in this work, we focus our attention on establishing the feasibility of SSC estimation using RSS.

Regarding low frequency wireless bands, a non-destructive method proposed in [21] took 10 to 1800 MHz signals to measure the dielectric properties and to detect the sugar content of three apple breeds. The Rician k-factor of 2.45 GHz signal response was used to estimate the granulated parts in tangerine [29]. Authors in [15] proposed to measure banana sugar content using ultrasound velocity, and similar to other approaches, such techniques require new dedicated hardware devices which are not readily available to consumers. Authors in [41] used mmWave signals to estimate sugar content in apples. However, the focus there is on a different mmWave frequency (40 GHz). Alternatively, our work uses 60 GHz frequency which is also an ISM frequency band. Also, compared to [41], we provide regression models that use RSS and can be readily deployed on consumer's smartphone equipped with off-the-shelf 60 GHz radios.

9 CONCLUSIONS

In this paper, we developed a soluble sugar estimation technique where 60 GHz mmWave signal can be reflected off fruits to estimate their Brix values. First, we showed how the sugar content changes the permittivity which in turn affect the RSS of the reflected signal. We then systematically studied different factors of fruits such as fruit diameter, flesh density and surface roughness to show that these factors have non-trivial impacts on reflection. This motivates us to develop separate per-fruit estimation model using regression. We derived two types of features (amplitude and RSS) and showed how they can be useful in Brix estimation. With extensive evaluation, we showed that our per fruit type regression models that use all features can achieve a mean correlation coefficient of 85% (apples: 86%, oranges: 85% and kiwis: 84%). We also developed a common model across all fruit types and found that due to the diversity of fruit characteristics such models perform poorly in SSC estimation. Lastly, we showed that when only RSS based features are used, the SSC estimation can reach the mean correlation coefficient of 78%. Given that our research focuses on establishing the feasibility of SSC estimation using mmWave, it opens up new directions of research and application development as discussed in the previous section.

REFERENCES

[1] Jorgen Bach Andersen, Theodore S Rappaport, and Susumu Yoshida. 1995. Propagation measurements and models for wireless communications channels. *IEEE Communications Magazine* (1995).
[2] S Arazuri, C Jarén, and JI Arana. 2005. Selection of the temperature in the sugar content determination of kiwi fruit. *International journal of infrared and millimeter waves* (2005).
[3] Philip Ashurst, Robert Hargitt, and Fiona Palmer. 2017. *Soft drink and fruit juice problems solved.* Woodhead Publishing.
[4] ATAGO. 2017. *Pocket Brix-Acidity Meter Multi Fruits PAL-BX|ACID F5 Master Kit.* http://www.atago.net/product/?l=ue&f=products_acid.php.

[5] Jie Bao and Moungi G Bawendi. 2015. A colloidal quantum dot spectrometer. *Nature* (2015).

[6] FJ Bates et al. 1942. Polarimetry, Saccharimetry and the Sugars, US Government Printing Office, Washington, DC.

[7] HE Bennett and JO Porteus. 1961. Relation between surface roughness and specular reflectance at normal incidence. *JOSA* (1961).

[8] J Temple Black and Ronald A Kohser. 2017. *DeGarmo's materials and processes in manufacturing*. John Wiley & Sons.

[9] Leo Breiman. 2001. Random forests. *Machine learning* (2001).

[10] Kenneth S Cole and Robert H Cole. 1941. Dispersion and absorption in dielectrics I. Alternating current characteristics. *The Journal of chemical physics* (1941).

[11] Jorge R Costa, Eduardo B Lima, and Carlos A Fernandes. 2009. Compact beam-steerable lens antenna for 60-GHz wireless communications. *IEEE Transactions on Antennas and Propagation* (2009).

[12] Caitlin Covington. 2014. *Sugar Wise: How Fruit Stacks Up.* https://greatist.com/health/sugar-wise-how-fruits-stack.

[13] Peter Josef William Debye. 1929. *Polar molecules.* Chemical Catalog Company, Incorporated.

[14] Deelat. 2016. *Digital Refractometer – Portable - D1160705.* https://www.deelat.com/tools/test-measurement-devices/refractometers/digital-refractometer-portable-d1160705.

[15] M Azfar Din, MH Abdul Halim, and Norlida Buniyamin. 2016. Determination of banana sugar content using ultrasound velocity measurement. In *Information Technology and Electrical Engineering (ICITEE), 2016 8th International Conference on.* IEEE, 1–5.

[16] Harris Drucker, Christopher JC Burges, Linda Kaufman, Alex J Smola, and Vladimir Vapnik. 1997. Support vector regression machines. In *Advances in neural information processing systems.*

[17] K Fuchs and U Kaatze. 2002. Dielectric spectra of mono-and disaccharide aqueous solutions. *The Journal of chemical physics* (2002).

[18] Edoardo Gatti, Nicola Di Virgilio, Massimiliano Magli, and Stefano Predieri. 2011. Integrating sensory analysis and hedonic evaluation for apple quality assessment. *Journal of Food Quality* (2011).

[19] Mayank Goel, Eric Whitmire, Alex Mariakakis, T Scott Saponas, Neel Joshi, et al. 2015. HyperCam: hyperspectral imaging for ubiquitous computing applications. In *ACM Ubicomp 2015.* ACM.

[20] David E. Goldberg. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley Longman Publishing Co., Inc.

[21] Wenchuan Guo, Liang Shang, Xinhua Zhu, and Stuart O Nelson. 2015. Nondestructive detection of soluble solids content of apples from dielectric spectra with ANN and chemometric methods. *Food and Bioprocess Technology* (2015).

[22] Mark Andrew Hall. 1999. Correlation-based feature selection for machine learning. (1999).

[23] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool release: Gathering 802.11 n traces with channel state information. *ACM SIGCOMM Computer Communication Review* (2011).

[24] FR Harker, KB Marsh, H Young, SH Murray, FA Gunson, and SB Walker. 2002. Sensory interpretation of instrumental measurements 2: sweet and acid taste of apple fruit. *Postharvest Biology and Technology* 24, 3 (2002), 241–250.

[25] IEEE. 2005. *Transmission thru layered structures at 60 GHz.* https://mentor.ieee.org/802.15/dcn/05/15-05-0388-01-003c-transmission-caculations-60ghz.doc.

[26] PD Jamieson, JR Porter, and DR Wilson. 1991. A test of the computer simulation model ARCWHEAT1 on wheat crops grown in New Zealand. *Field crops research* (1991).

[27] Asutosh Kar and Leena Das. 2011. A technical review on statistical feature extraction of ecg signal. In *IJCA Special Issue on 2nd National Conference-Computing, Communication and Sensor Network, CCSN.*

[28] Dan Kimball. 2012. *Citrus processing: a complete guide.* Springer Science & Business Media.

[29] Chainarong Kittiyanpunya, Prapan Leekul, Chuwong Phongcharoenpanich, and Monai Krairiksh. 2017. Beam-Scanning Reflectometer for Detecting Granulated Fruits. *IEEE Sensors Journal 2017* (2017).

[30] Tryphon Lambrou, Panos Kudumakis, R Speller, M Sandler, and A Linney. 1998. Classification of audio signals using statistical features on time and wavelet transform domains. In *IEEE ICASSP 1998.* IEEE.

[31] Laxco. 2017. *Handheld Digital Brix/RI Refractometer.* https://docs.google.com/a/laxcoinc.com/viewer?a=v&pid=sites&srcid=bGF4Y29pbmMuY29tfHd3d3xneDo2NzBhNzRjMjU1NjUyYzZl.

[32] A Lerbret, Patrice Bordat, F Affouard, M Descamps, and F Migliardo. 2005. How homogeneous are the trehalose, maltose, and sucrose water solutions? An insight from molecular dynamics simulations. *The Journal of Physical Chemistry B* (2005).

[33] Roger Lhermitte. 1990. Attenuation and scattering of millimeter wavelength radiation by clouds and precipitation. *Journal of Atmospheric and Oceanic Technology* 7, 3 (1990), 464–479.

[34] Renfu Lu. 2001. Predicting firmness and sugar content of sweet cherries using near–infrared diffuse reflectance spectroscopy. *Transactions of the ASAE* 44, 5 (2001), 1265.

[35] Chunsheng Luo, Long Xue, Muhua Liu, Jing Li, and Xiao Wang. 2010. Non-destructive measurement of sugar content in navel orange based on Vis-NIR spectroscopy. In *Springer CCTA 2010.*

[36] Gustav Mie. 1908. Beiträge zur Optik trüber Medien, speziell kolloidaler Metallö-sungen. *Annalen der physik* 330, 3 (1908), 377–445.

[37] David S Moore, George P McCabe, and Bruce A Craig. 2009. *Introduction to the Practice of Statistics.* WH Freeman New York.

[38] Bart M Nicolai, Karen I Theron, and Jeroen Lammertyn. 2007. Kernel PLS regression on wavelet transformed NIR spectra for prediction of sugar content of apple. *Chemometrics and intelligent laboratory systems* (2007).

[39] Thomas Nitsche, Adriana B Flores, Edward W Knightly, and Joerg Widmer. 2015. Steering with eyes closed: Mm-Wave beam steering without in-band measurement. In *Computer Communications (INFOCOM), 2015 IEEE Conference on.* IEEE, 2416–2424.

[40] Hyun Kwon Noh and Renfu Lu. 2007. Hyperspectral laser-induced fluorescence imaging for assessing apple fruit quality. *Postharvest Biology and Technology* 43, 2 (2007), 193–201.

[41] Makoto Oda, Atsushi Mase, Kiichiro Uchino, et al. 2011. Non-destructive measurement of sugar content in apples using millimeter wave reflectometry and artificial neural networks for calibration. In *Asia-Pacific Microwave Conference (APMC 2011): 25 Years, The.* Engineers Australia, 1386.

[42] Ahmad Fairuz Omar, Hanafi Atan, and Mohd Zubir MatJafri. 2012. Peak response identification through near-infrared spectroscopy analysis on aqueous sucrose, glucose, and fructose solution. *Spectroscopy Letters* (2012).

[43] ITU-R P.2040-1. 2015. *Effects of building materials and structures on radiowave propagation above about 100 MHz.* International Telecommunication Union Recommendation ITU-R.

[44] John J Palmer. 2006. *How to Brew: Everything you need to know to brew beer right the first time.* Brewers Publications.

[45] Shimul Saha, Helena Cano-Garcia, Ioannis Sotiriou, et al. 2017. A Glucose Sensing System Based on Transmission Measurements at Millimetre Waves using Micro strip Patch Antennas. *Scientific Reports* (2017).

[46] Katsuyoshi Sato, Hideki Kozima, Hiroshi Masuzawa, et al. 1995. Measurements of reflection characteristics and refractive indices of interior construction materials in millimeter-wave bands. In *Vehicular Technology Conference, 1995 IEEE 45th.*

[47] Estanislao Silla, Arturo Arnau, and Iñaki TuñóN. 2001. Fundamental principles governing solvents use. *Handbook of Solvents* (2001), 7.

[48] DC Slaughter and JA Abbott. 2004. 14 Applications in Analysis of Fruits and Vegetables. (2004).

[49] Sanjib Sur, Vignesh Venkateswaran, Xinyu Zhang, and Parmesh Ramanathan. 2015. 60 GHz Indoor Networking Through Flexible Beams: A Link-Level Profiling. In *Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (ACM SIGMETRICS '15).*

[50] Deepak Vasisht, Zerina Kapetanovic, Jongho Won, et al. 2017. FarmBeats: An IoT Platform for Data-Driven Agriculture.. In *NSDI.* 515–529.

[51] VubIQ. 2017. *60 GHz Development System.* https://www.pasternack.com/.

[52] Robert C Weast. 1972. *Handbook of Chemistry and Physics 53rd Edition.* Chemical Rubber Pub.

[53] Teng Wei and Xinyu Zhang. 2015. mtrack: High-precision passive tracking using millimeter wave radios. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking.* ACM, 117–129.

[54] Wikipedia. 2017. *Brix.* https://en.wikipedia.org/wiki/Brix.

[55] Xin-She Yang. 2010. Firefly algorithm, stochastic test functions and design optimisation. *International Journal of Bio-Inspired Computation* (2010).

[56] Zhicheng Yang, Parth H Pathak, Jianli Pan, Mo Sha, and Prasant Mohapatra. 2018. Sense and Deploy: Blockage-aware Deployment of Reliable 60 GHz mmWave WLANs. In *2018 IEEE MASS.* IEEE.

[57] Zhicheng Yang, Parth H Pathak, Yunze Zeng, Xixi Liran, and Prasant Mohapatra. 2017. Vital sign and sleep monitoring using millimeter wave. *ACM TOSN* (2017).

[58] Zhicheng Yang, Parth H Pathak, Yunze Zeng, and Prasant Mohapatra. 2015. Sensor-assisted codebook-based beamforming for mobility management in 60 ghz wlans. In *IEEE MASS 2015.* IEEE, 333–341.

[59] Turker Yilmaz, Etimad Fadel, and Ozgur B Akan. 2014. Employing 60 GHz ISM band for 5G wireless communications. In *IEEE BlackSeaCom 2014.* IEEE.

[60] Huanle Zhang, Ahmed Elmokashfi, Zhicheng Yang, and Prasant Mohapatra. 2018. Wireless Access to Ultimate Virtual Reality 360-Degree Video At Home. *arXiv preprint arXiv:1812.01777* (2018).

[61] Xia Zhou, Zengbin Zhang, Yibo Zhu, Yubo Li, Saipriya Kumar, Amin Vahdat, Ben Y Zhao, and Haitao Zheng. 2012. Mirror mirror on the ceiling: flexible wireless links for data centers. *ACM SIGCOMM Computer Communication Review.*

[62] Yanzi Zhu, Yuanshun Yao, Ben Y Zhao, and Haitao Zheng. 2017. Object recognition and navigation using a single networking device. In *ACM MobiSys 2017.*

[63] Yanzi Zhu, Yibo Zhu, Ben Y Zhao, and Haitao Zheng. 2015. Reusing 60ghz radios for mobile radar imaging. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking.* ACM, 103–116.

[64] Bruce W Zoecklein, Kenneth C Fugelsang, Barry H Gump, and Fred S Nury. 1995. *Wine analysis and production.* Springer.