

Characterization of Range for Smart Home Sensors Using Tsallis' Entropy Framework



Sujit Bebortta, Amit Kumar Singh, Surajit Mohanty and Dilip Senapati

Abstract The deployment of sensor nodes (SNs) in smart homes induces multipath transmission of signals in indoor environments (IEs). These paths occur due to the presence of household utility objects, which produce several reflected communication paths between the sender and the receiver. In order to determine the effective position of a target SN in a wireless sensor network (WSN), certain localization schemes are required in conjunction with trilateration methods. Therefore, it is worthwhile to estimate the uncertainties in the range for time of arrival (TOA) localization of SNs subjected to non-line of sight (NLOS) conditions. In this article, we provide a technique to characterize the variations in the range corresponding to the TOA based on the well-known Tsallis' entropy framework. In this model, the non-extensive parameter q characterizes the variations in the localization range caused due to multipath components. In this context, we optimize the Tsallis entropy subject to the two moment constraints (i.e., mean and variance) along with the normalization constraint. Our proposed model is in excellent agreement with the synthetic data in contrast to the mixture model. This paper also provides a new approach for estimating the parameters corresponding to the mixture model and the proposed q -Gaussian model by minimizing the Jensen-Shannon (JS) symmetric measure between the two models and the synthetic data.

Keywords q -Gaussian distribution · Tsallis entropy · Mixture model · Wireless sensor networks · Localization · JS measure

S. Bebortta · D. Senapati (✉)

Department of Computer Science Ravenshaw University, Cuttack 753003, India
e-mail: senapatidillip@gmail.com

A. K. Singh

Department of Computer Science Ramanujan College, University of Delhi,
New Delhi 110019, India

S. Mohanty

Department of Computer Science and Engineering DRIEMS Tangi,
Cuttack, Orissa 754022, India

© Springer Nature Singapore Pte Ltd. 2020

B. Pati et al. (eds.), *Advanced Computing and Intelligent Engineering*,
Advances in Intelligent Systems and Computing 1089,
https://doi.org/10.1007/978-981-15-1483-8_23

1 Introduction

In the recent times, the development of the smart homes has drastically increased with the availability of various smart sensors and energy-efficient resources. Several smart devices like the smartphones, televisions, HVAC (heating, ventilation, and air conditioning) systems, etc., have contributed to the growth of these smart homes. These technologies employ several sensors for acquiring data from varying locations. So, it is quite difficult to infer the actual location of a target sensor node (SN) due to the ubiquity of scattering objects and multiple transmission paths. Thus, for transmission in indoor environments (IEs), the ultra-wideband (UWB) technology is considered to be the most reliable approach for facilitating communication between the smart devices [1–4, 26]. The UWB technology has marked a prime development in wireless communication industries, with the ability to support high data rate transmission over a short range. It also consumes less power which makes it an energy-efficient choice for personal area networks (PANs). The time of arrival (TOA)-based ranging is used distinctively with the UWB technology for localizing the sensors in a wireless sensor network (WSN). The TOA ranging can make complete utilization of the high bandwidth and appropriately manages the time delay resolution of the UWB technology [14, 15].

In IEs, the data transmission path between the source and the sink can be located either by using a line of sight (LOS) or non-line of sight (NLOS) propagation. In smart homes, the NLOS propagation indicates a lack of direct communication path between the source and the sink, due to the presence of obstacles, *viz.*, furniture and walls. A major concern with indoor complex networks is the transmission of data over the UWB channel in NLOS environments. The execution of a TOA-based UWB communication relies on accessibility of a direct signaling path between the source and the sink [10]. Thus due to the ambiguity in the existence of an explicit data path between the source and the sink, the TOA localization can exhibit significant variations in its range.

In this paper, we employ a straightforward approach to characterize the variations in range by using the well-known Tsallis' entropy framework. This framework was proposed to depict the consolidation of the power-law behavior and tail fluctuations of dynamical systems in statistical mechanics [5–9]. We first consider the localization of SNs by using the UWB-based TOA approach to infer the location of an SN. In IEs, the uncertainties in range caused due to multiple paths can be characterized by using the proposed model, i.e., the q -Gaussian distribution which is determined by maximizing the Tsallis entropy. Here, we consider a network of sensor nodes supporting multipath transmission; thus, a prior information about the location of a set of SNs is essential in order to infer the location of other SNs in the network. Therefore, by using an efficient sensor localization techniques, the initial cost associated with the deployment of SNs, as well as the cost for powering each node in a WSN, can be reduced greatly. As we are considering a multipath UWB communication between the SNs, the prime source of error is due to the presence of multiple paths between the transmitter and the receiver. Therefore, due to the presence of these errors originating

from various points in a WSN, the estimates obtained using the TOA localization get degraded [10, 16]. So, it becomes quite challenging and inevitable to characterize the variations in range for multipath communication channels. Here, we employ maximum Tsallis' entropy framework subjected to the normalization constraint, and the first two moment constraints, which yields the q -Gaussian distribution. The different values of the non-extensive parameter q characterize the variations in range. Beck and Cohen [18] proposed a superstatistics framework for modeling complex systems, by averaging over the environmental fluctuations. Thus by following [18], we design a mixture model for the variations in range by incorporating the fluctuations caused due to multipath components.

We also present a robust technique for estimation of the parameters corresponding to the PDFs of the mixture model and the proposed model, in order to characterize the synthetic data.

The rest of this paper is as follows: Sect. 2 illustrates the problem statement. In Sect. 3, we derive the q -Gaussian distribution using the maximum Tsallis' entropy framework. The parameter estimation techniques corresponding to the mixture model and our proposed model are discussed in Sect. 4. Section 5 deals with the characterization of range for different values of q . Finally, in Sect. 6 we provide the conclusion and future works.

2 Problem Statement

When a signal travels from the sender (or transmitter) to a receiver, then its propagation time can be estimated by using the TOA technique. Generally, TOA estimation is used along with UWB technology for an energy-efficient ranging in IEs with multiple propagation paths. We use localization of SNs to quantify the signal obtained from a heterogeneous set of sensors. Figure 1 illustrates the deployment of several SNs in a smart home. These SNs are positioned abundantly throughout the smart home and can be used to sense diverse indoor environments.

Figure 2 shows the existence of multiple transmission paths in ubiquity of several interfering objects. These signals undergo scattering and reflection due to interference of several objects present in an indoor environment. The interfering objects scatter the waves while being transmitted from the transmitter (Tx) to the receiver (Rx) which leads to difficulty in identifying the actual location of a target node.

We can model the UWB channels having multipath components by utilizing the impulse response [11], i.e.,

$$h(t) = \sum_{l=0}^m a_l \delta(t - \tau_l), \quad (1)$$

where a_l is the amplification factor for the l th multipaths and τ_l represents the time delay of the channel.



Fig. 1 Deployment of multiple sensors in a smart home environment

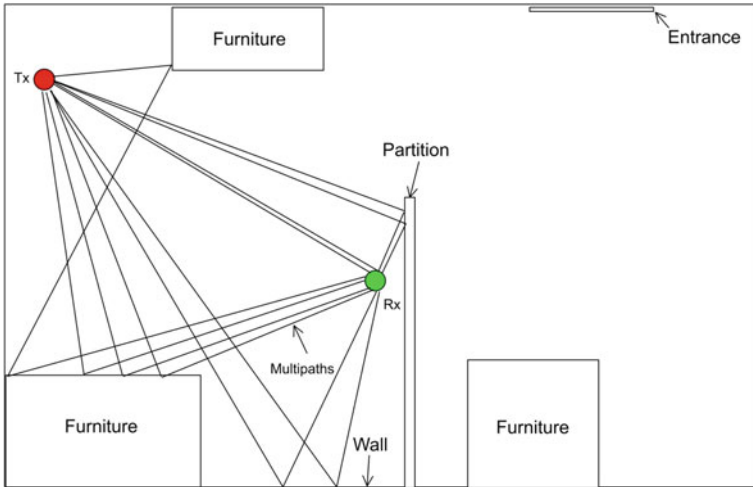


Fig. 2 Variations in range under the presence of multiple paths in IEs

Therefore, using Eq. (1), the received signal corresponding to the TOA estimate can be characterized using [13] as

$$r_{i,l}(t) = \sum_{i=0}^n \sum_{l=0}^m a_{i,l} s(t - \tau_{i,l}) + n_{i,l}(t), \tag{2}$$

where $s(t)$ represents the transmitted signal and $n_{i,l}(t)$ denotes the additive white Gaussian noise. Thus, it can be observed that under NLOS conditions these channels

reflect dense multipath effects and time delays which may affect the overall performance of the system.

Let us consider the sensor nodes to be deployed in an indoor environment. If we consider the TOA ranging for determining the location of the sensor nodes, then their range can be given as

$$r_i = d_i + n_i, \quad i = 0, 1, 2, \dots, n, \tag{3}$$

where r_i is the measure of the range between a source and receiver for i th sensor nodes and d_i is the distance between the source and the i th sensor and can be stated as

$$d_i = \|x - x_i\|_2 = \sqrt{(x - x_i)^2 + (y - y_i)^2}, \tag{4}$$

here, $(x, y) \in X$ is a set of remote source locations and $(x_i, y_i) \in X_i$ is the set of recognized coordinates of the i th sensor, and n_i is the additive white Gaussian noise, such that $n_i \sim N(0, \sigma_i^2)$.

Thus if we characterize the range in terms of a Gaussian distribution [19], then its probability density function (PDF) can be expressed as

$$f(r_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(r_i - d_i)^2}{2\sigma_i^2}}, \tag{5}$$

where $r_i \sim N(d_i, \sigma_i^2)$, with σ_i^2 representing the variance corresponding to r_i . However, the variations (σ_i^2) in the range r_i can be characterized by inverse Gamma distribution [17, 18, 20]. So, the parameter σ_i can be considered to be a random variable such that

$$\beta_i = \frac{1}{\sigma_i^2} \sim \text{Gamma}(a, b), \tag{6}$$

where a and b are the shape and scale parameters, and β_i is the inverse of the variations caused in the range for i th SNs.

Let $g(\sigma_i)$ be the variational range distribution and be defined as

$$g(\sigma_i) = \frac{1}{b^a \Gamma(a)} \beta_i^{a-1} e^{-\frac{\beta_i}{b}}, \quad \beta_i \geq 0. \tag{7}$$

The PDF for the gamma distribution in compliance with the synthetic data is illustrated in Fig. 5.

The unconditional distribution of range is obtained by averaging over the gamma distribution [17, 25],

$$p(r_i) = \int_0^\infty f(r_i|\sigma_i) g(\sigma_i) d\sigma_i, \tag{8}$$

where $f(r_i|\sigma_i)$ is the conditional probability distribution of r_i for a given σ_i . Now, we have

$$p(r_i) = \int_0^\infty \sqrt{\frac{\beta_i}{2\pi}} e^{-\frac{\beta_i(r_i-d_i)^2}{2}} \frac{1}{b^a \Gamma(a)} \beta_i^{a-1} e^{-\frac{\beta_i}{b}} \frac{1}{2} (\beta_i)^{-\frac{3}{2}} d\beta_i. \tag{9}$$

So, the probability distribution corresponding to the variations in range observed in an indoor multipath environment is given as

$$p(r_i) = \frac{\Gamma(a-1)}{2\sqrt{2\pi}b^a \Gamma(a)} \left[\frac{1}{b} + \frac{(r_i-d_i)^2}{2} \right]^{-(a-1)}. \tag{10}$$

The above mixture model along with the synthetic data is shown in Fig. 6.

3 Proposed Model

In order to obtain accuracy in inferring the location of SNs in an IE, we need to characterize the variations in the measurement of the range observed due to multipath components. If we consider r_i to be a random variable, then by applying Tsallis entropy [6, 7, 21–23], the uncertainties in the range obtained in Eq. (3) can be characterized through the non-extensive parameter q and is expressed as

$$S_q(r_i) = \frac{1 - \int_{-\infty}^\infty [f_q(r_i)]^q dr_i}{q-1}. \tag{11}$$

Here, we investigate the probability distribution corresponding to the variations in range for smart home sensors by maximizing the Tsallis entropy framework subjected to the normalization constraint and the moment constraints.

Thus, the constraint for normalization can be given as

$$\int_{-\infty}^\infty f_q(r_i) dr_i = 1. \tag{12}$$

The first two constraints for the moments of r_i can be specified as

$$\int_{-\infty}^\infty f_{es}(r_i) r_i dr_i = \mu, \tag{13}$$

and

$$\int_{-\infty}^{\infty} f_{es}(r_i - \mu)^2 dr_i = \sigma^2, \tag{14}$$

where μ and σ^2 indicate the mean and variance.

Therefore, the probability distribution in Eqs. (13) and (14) corresponding to any probability f_q can be expressed as

$$f_{es}(r_i) = [f_q(r_i)]^q / \int [f_q(r_i)]^q dr_i, \tag{15}$$

where $f_{es}(r_i)$ denotes the escort probability and is used to represent the asymptotic decay of q moments. Since we are using a parameter q , so the type of expectation used for defining the variations in the range is called the q -expectation. Thus, by considering the Lagrangian, for any stationary SN, we can optimize the range of the received signal as

$$\begin{aligned} L(f_q, r_i) = & \frac{1 - \int_{-\infty}^{\infty} f_q(r_i)^q dr_i}{(q - 1)} + \omega_1 \left(1 - \int_{-\infty}^{\infty} f_q(r_i) dr_i \right) \\ & + \omega_2 \left(\mu - \int_{-\infty}^{\infty} r_i f_{es}(r_i) d_i \right) + \omega_3 \left(\sigma^2 - \int_{-\infty}^{\infty} f_{es}(r_i - \mu)^2 dr_i \right), \end{aligned} \tag{16}$$

where ω_1, ω_2 and ω_3 are the Lagrangian parameters. Now by employing the Euler-Lagrange equation, we have

$$\frac{\partial L}{\partial f_q(r_i)} - \frac{d}{dr_i} \left(\frac{\partial L}{\partial f_q'(r_i)} \right) = 0 \tag{17}$$

Thus, by using Eq. (17) in association with Eqs. (11)–(16), we have

$$f_q(r_i) = \frac{1}{z} \left(1 - \frac{1 - q}{(3 - q)\sigma^2} (r_i - \mu)^2 \right)^{1/(1 - q)}, \quad 1 < q < 3, \tag{18}$$

where $z = \sigma \sqrt{\frac{(3 - q)\pi}{q - 1} \frac{\Gamma(\frac{3 - q}{2q - 2})}{\Gamma(\frac{1}{q - 1})}}$ and is called the normalization constant. From Eq. (18), we obtain the maximum value for the entropy probability distribution corresponding to the variations in range of the TOA estimate.

4 Parameter Estimation

In this section, we discuss the parameter estimation for the mixture model obtained in Eq. (10) and the proposed q -Gaussian distribution obtained in Eq. (18). The shape parameter ‘a’ and scale parameter ‘b’ of the mixture distribution can be estimated through the probability distribution of ‘ σ_i ’. The random variable σ_i is defined as

$$\sigma_i (r_i) \equiv \left[\frac{1}{N} \sum_{\tau=1}^N |v_\tau [t : t + \tau]| \right]^2, \quad i = 1, 2, \dots, n, \tag{19}$$

where N is the number of samples per time stamp (τ) and v_τ represents the variations in the range within the interval $[t, t + \tau]$.

The estimate of non-extensive parameter ‘ q ’ can be expressed through the generalized Jensen–Shannon (JS) symmetric measure. Thus, the symmetric JS divergence measure can be defined as

$$\begin{aligned} \text{JS} (f_q (r_i) || g (r_i)) &= \frac{1}{2} \text{KL} \left(f_q (r_i) || \frac{f_q (r_i) + g (r_i)}{2} \right) \\ &+ \frac{1}{2} \text{KL} \left(g (r_i) || \frac{f_q (r_i) + g (r_i)}{2} \right) \end{aligned} \tag{20}$$

where $KL (.)$ is the generalized Kullback–Leibler measure. This measure provides the distance between the proposed distribution $f_q (r_i)$ and the empirical distribution $g (r_i)$ [7, 24]. Thus, the KL measure is given as

$$\text{KL} (f_q (r_i) || g (r_i)) = \sum_{k=1}^n f_{q_k} (r_i) \log_2 \left| \frac{f_q (r_i)}{g_k (r_i)} \right| \tag{21}$$

The optimal value corresponding to the non-extensive parameter q is attained at which the symmetric JS divergence is minimum.

Thus, the optimal value of the \hat{q} estimate provides a closed agreement between the synthetic data and the proposed theoretical distribution. Therefore, the optimal value of q is estimated as

$$\hat{q} = \arg \min_{q \in (1,3)} \{ \text{JS} (f_q (r_i) || g (r_i)) \} \tag{22}$$

5 Results and Discussion

In this section, we plot the PDF for the maximum entropy obtained in Sect. 3 for the different values of q , viz., $q = 1.5$ (proposed model), $q = 1.37$ and $q = 1.45$. It can be observed that at $q = 1.5$, our model successfully captures the variations in range

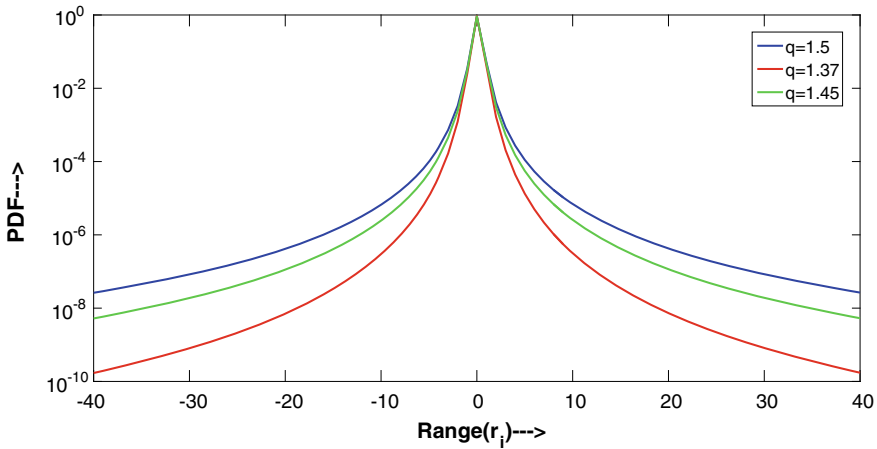


Fig. 3 q -Gaussian PDFs for different values of q , viz., $q = 1.5$ (proposed model), $q = 1.37$ and $q = 1.45$

caused due to indoor multipath propagation. It is worth noting that for $q = 1.5$ (as shown in Fig. 3), the model exhibits the long memory behavior and extensively entails the extreme fluctuations in the range. In order to validate our model, we generate synthetic data using MATLAB incorporating multipath components. Figure 4a shows the trend plot for the synthetic data with a sample size of 1000, and Fig. 4b shows its respective histogram. We transformed the variations in range for indoor environments to the gamma distribution, obtained in Eq. (7), and its compliance with the synthetic data is shown in Fig. 5. In IEs, the presence of dense multipath and the ubiquity of NLOS conditions causes degradation in the estimation of TOA. From Fig. 6, it can be observed that our proposed framework captures the dense multipath components of the generated signal even at the tail end. From the above discussions, we can draw an inference that when $q = 1.5$, the model provides a better characterization for the uncertainties in the TOA estimation. Further, we can obtain greater precision for estimating the range of SNs, using TOA with resolutions for providing more accuracy in temporal and multipath components.

6 Conclusion and Future Work

In this paper, we discussed the UWB-based TOA localization and variations in its range for indoor environments. We characterized the variations in range for TOA localization by using Tsallis' entropy framework. The maximized Tsallis entropy subjected to the first two moment constraints (i.e., mean and variance) along with the normalization constraint is derived. It is observed that our model captures the variations in the localization range in the existence of dense multipaths in indoor

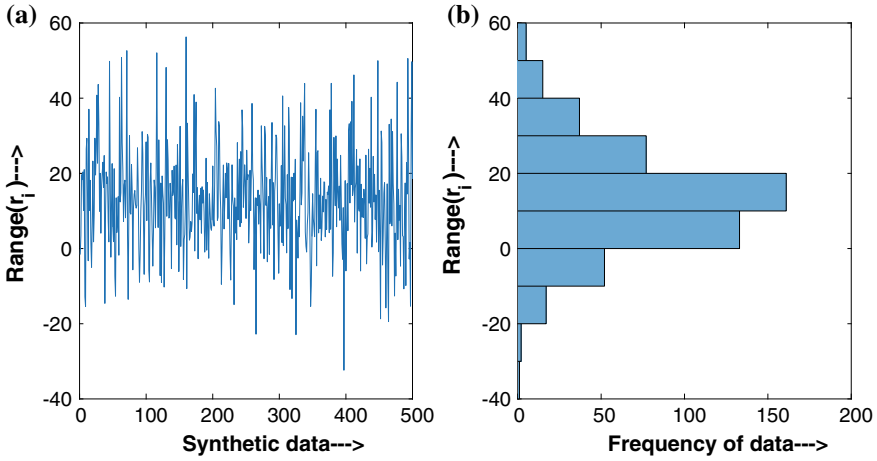


Fig. 4 a Representation of the trend plot with sample size 1000. b Representation of histogram corresponding to the trend plot

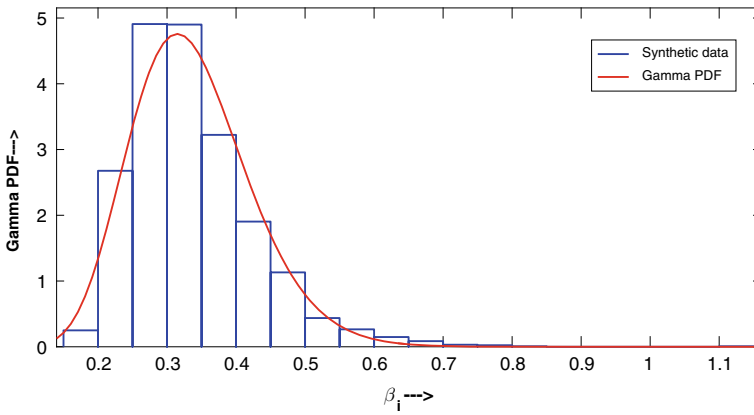


Fig. 5 Representation of the gamma PDF in convergence with the synthetic data for $a = 9.04432$ and $b = 0.0285463$

environments. Our proposed q -Gaussian model operates well over the synthetic data in contrast to the mixture model. It is observed that the q -Gaussian distribution provided a better characterization of the variations in the localization range for $q = 1.5$. This paper also presented a new approach for estimation of parameters corresponding to the mixture model and our proposed model, by minimizing the JS symmetric measure between the two models and the synthetic data.

In the future, we will provide an intrinsic stochastic differential equation (SDE) model and its corresponding Fokker-Planck equation, to entail the variations in the localization range for densely positioned indoor environments.

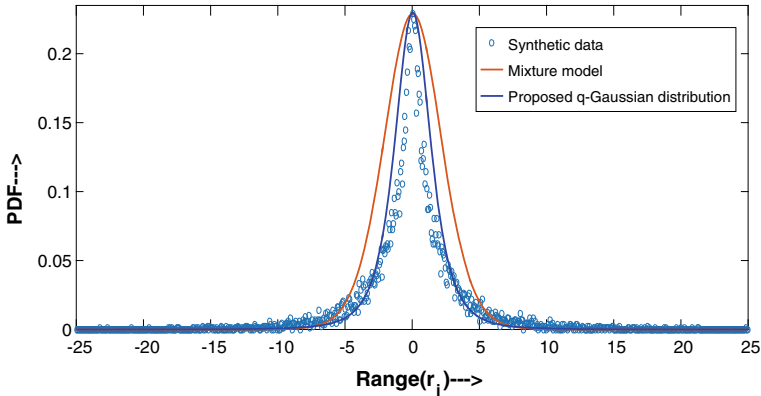


Fig. 6 Fitting of PDFs for the mixture model for $a = 9.04432$ and $b = 0.0285463$ along with the proposed q -Gaussian model over the synthetic data for uncertainties in range, obtained for $q = 1.5$ with $\mu = 0.0428$ and $\sigma = 1.4138$.

References

1. Porcino, D., Hirt, W.: Ultra-wideband radio technology: potential and challenges ahead. *IEEE Commun. Mag.* **41**(7), 66–74 (2003)
2. Oppermann, I., Hämäläinen, M., Iinatti, J. (eds.): *UWB: Theory and Applications*. Wiley & Sons (2005)
3. Rajput, N., Gandhi, N., Saxena, L.: Wireless sensor networks: apple farming in northern India. In: 2012 Fourth International Conference on Computational Intelligence and Communication Networks (CICN). IEEE (2012)
4. Gezici, S., et al.: Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks. *IEEE Signal Process. Mag.* **22**(4), 70–84 (2005)
5. Tsallis, C., et al.: Nonextensive statistical mechanics and economics. *Physica A Stat. Mech. Appl.* **324**(1–2), 89–100 (2003)
6. Tsallis, C., Non-extensive Statistical Mechanics: Construction and physical interpretation. In: *Nonextensive Entropy Interdisciplinary Applications*, pp. 1–52 (2004)
7. Senapati, D., Karmeshu: Generation of cubic power-law for high frequency intra-day returns: maximum Tsallis entropy framework. *Digit. Signal Process.* **48**, 276–284 (2016)
8. Singh, A.K.: Karmeshu: Power law behavior of queue size: maximum entropy principle with shifted geometric mean constraint. *IEEE Commun. Lett.* **18**(8), 1335–1338 (2014)
9. Singh, A.K., Singh, H.P., Karmeshu: Analysis of finite buffer queue: maximum entropy probability distribution with shifted fractional geometric and arithmetic means. *IEEE Commun. Lett.* **19**(2), 163–166 (2015)
10. Lee, J.-Y., Scholtz, R.A.: Ranging in a dense multipath environment using an UWB radio link. *IEEE J. Sel. Areas Commun.* **20**(9), 1677–1683 (2002)
11. Ghavami, M., Michael, L., Kohno, R.: *Ultra wideband signals and systems in communication engineering*. Wiley & Sons (2007)
12. Abe, S., Bagci, G.B.: Necessity of q -expectation value in nonextensive statistical mechanics. *Phys. Rev. E* **71**(1), 016139 (2005)
13. Saleh, A.A.M., Valenzuela, R.: A statistical model for indoor multipath propagation. *IEEE J. Sel. Areas Commun.* **5**(2), 128–137 (1987)
14. Pietrzyk, M.M., von der Grün, T.: Experimental validation of a TOA UWB ranging platform with the energy detection receiver. 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE (2010)

15. Alavi, B., Pahlavan, K.: Modeling of the TOA-based distance measurement error using UWB indoor radio measurements. *IEEE Commun. Lett.* **10**(4), 275–277 (2006)
16. Zhou, Y., et al.: Indoor elliptical localization based on asynchronous UWB range measurement. *IEEE Trans. Instrum. Meas.* **60**(1), 248–257 (2011)
17. Gerig, A., Vicente, J., Fuentes, M.A.: Model for non-Gaussian intraday stock returns. *Phys. Rev. E* **80**(6), 065102 (2009)
18. Beck, C., Cohen, E.G.D.: Superstatistics. *Physica A Stat. Mech. Appl.* **322**, 267–275 (2003)
19. Zhao, W., et al.: Outage analysis of ambient backscatter communication systems. *IEEE Commun. Lett.* **22**(8), 1736–1739 (2018)
20. Beck, C., Cohen, E.G.D., Swinney, H.L.: From time series to superstatistics. *Phys. Rev. E* **72**(5), 056133 (2005)
21. Rajput, N.K., Ahuja, B., Riyal, M.K.: A novel approach towards deriving vocabulary quotient. *Digit. Sch. Humanit.* **33**, 894–901 (2018)
22. Riyal, M.K., et al.: Rank-frequency analysis of characters in Garhwali text: emergence of Zipf’s law. *Curr. Sci.* **110**(3), 429 (2016)
23. Schäfer, B., et al.: Non-Gaussian power grid frequency fluctuations characterized by Lévy-stable laws and superstatistics. *Nat Energy* **3**(2), 119 (2018)
24. Osán, T.M., Bussandri, D.G., Lamberti, P.W.: Monoparametric family of metrics derived from classical JensenShannon divergence. *Phys. A Stat. Mech. Appl.* **495**, 336–344 (2018)
25. Papadimitriou, K., Sfikas, G., Nikou, C.: Tomographic image reconstruction with a spatially varying gamma mixture prior. *J. Math. Imaging Vis.* 1–11 (2018)
26. Kaiwartya, O., Abdullah, A.H., Cao, Y., Rao, R.S., Kumar, S., Lobiyal, D.K., Isnin, I.F., Liu, X., Shah, R.R.: T-MQM: testbed-based multi-metric quality measurement of sensor deployment for precision agriculture—a case study. *IEEE Sensors J.* **16**(23), 8649–8664 (2016)