System design approach for heartbeat detection and classification of individuals irrespective of their physical condition

Chinmay Chandrakar* and Monisha Sharma

Department of Electronics and Telecommunication,
Shri Shankaracharaya College of Engineering and Technology,
Chhattisgarh Swami Vivekananda Technical University,
Bhilai 490 020, India

In an electrocardiogram (ECG), the heartbeat feature QRS is an important parameter for analysis in any heartbeat classification automated diagnosis system. In this communication the method which we have proposed is by using the counter which is used in parallel. The first level is detection of heartbeats, which uses hashing of ECG features. In the second level, the profiler profiles a person's regular and irregular ECG characteristic behaviour. The proposed method works on data related with ECG, instead of particular features of ECG. Because of parallel processing, data storage unit requirements and the processing time are less. The dependent values in the proposed method vary according to the changes in the ECG waveform. Such type of analysis is suitable for detection of heart disease. The most significant application of such characteristic plotting is to generate an alert signal for irregular ECG behaviour in a person. Such automated system will be useful in remote areas where a cardiologist may not be easily available.

Keywords: Data storage units, electrocardiogram signal, parallel processing, QRS detection.

THE determination and separation of QRS waveform between regular and irregular waveforms are particularly important clinical criteria for patient diagnosis. Over the last few decades, several techniques have been proposed to determine these waveforms[1–3]. For instance, Senhadji *et al.*[4] compared the performance of three different wavelets to determine the beats which are hidden in the waveform. In this paper, we propose a method for ECG waveform separation by using derivative of lower order in which Gaussian function was used[5]. Many more methods have been proposed which concentrated on spectral[6,7] or wavelet features[8,9], related with superimposed features[10,11] and spatial context[12,13] and to differentiate the heartbeats signal[14–19]. Pan and Tompkins[14] first proposed the method of determination of heartbeat in real time in processor. Ning and Selesnick[20] proposed that the location of the true peak can be determined which has the largest magnitude within its 200 ms time window. The ECG feature extractor provided by LabVIEW Biomedical toolkit detects QRS waves[21]. It was reported that for QRS determination, valley points before and after *R*-wave are sufficient[22]. It was also reported that the preprocessed ECG signal is converted into a train of pulses using the IF sampler[23]. Once the process of beat determination is complete, the correct beat discrimination process of heartbeats takes place. Methods such as detection of any symbol, repeated pattern which is going to repeat, network based on neurons, and vector machines have been suggested for heartbeat discrimination[24–26]. Learning methods based on training data from the sample which result in new mapping based on information which is processed has been discussed earlier[27]. de Chazal *et al.*[28] discriminated the beats by studying the distance between the two *R*-peaks, ECG waveform. Christov *et al.*[29] proposed comparison of ECG features for beat classification in time and frequency domain. Haseena *et al.*[30] discussed a combination of neuron and fuzzy for discrimination of heartbeats.

We have used arrhythmia database of the Massachusetts Institute of Technology–Beth Israel Hospital (MIT–BIH) to test the efficiency of the proposed technique[3]. All ECG data used here are sampled at 360 Hz, and the resolution of each is 8 bits/sample, therefore the bit rate of these data is 2880 bps. The method was run for both regular and irregular heartbeats.

High-frequency noise was removed using the simple time constant equation in which the window is moved from one *R*-peak to the next, while the low-frequency noise was removed using the frequency-domain transformation[31].

The proposed method depends more on a series of data of ECG waveform rather than any specific feature of the ECG waveform. The basic concept is to partition the input ECG signals into series of 0–1 strings. The next step is to select a string of *L* bytes. The start and end limits of this binary information are selected so that when long binary strings are repeated, they should lie on the same start and end limits. For this, the start and end limits are selected based on *R*-peak. Between these two repeated start and end points, the minimum value of binary data bits *Q* and *S* points starting from the middle *R*-peak to the left and right are determined for repeated occurrence within a fixed time duration. Hereafter, the data bits between *Q* and *S* points will be called as signature which is *W* bytes. *L* is varying as it depends on the start and end points of the *R*-peak, which basically works on *R*–*R* interval. Similarly, *W* is not fixed because it depends on the location of *Q* and *S* points. For the process to work in real-time QRS detection and classification, we used the storage elements which are arranged in parallel. Therefore in order to avoid the repeated number of counting in a memory elements string was hashed to a certain value. Two steps of hashing have been done in two phases. Phase 1 hash was first generated by the string. This hash

*For correspondence. (e-mail: chinmay_sscet@yahoo.co.in)

value was used as a signature for processing. The m independent hash values which have to be indexed into m arrays have been developed by this signature. Considering the memory address of the storage element is n bits, it means that each storage element is having 2^n memory locations. Each indexed array location is incremented. During the process, parallel counters are used. The counter which crosses the fixed value of count is detected as a QRS waveform. The cardiac profile is plotted against time duration of the suspected QRS waveform (in the form of number of samples) and the number of counts the suspected QRS waveform has repeated.

In phase 1, the window length is *R–R* interval, as shown in Figure 1. Every *R–R* interval length is processed. On every cycle of clock, one *R–R* interval string is shifted and a new byte for the packet stream of two consecutive *R–R* intervals is taken to form a next packet stream, i.e. boundaries (in bytes). From this *R*-peak, the *Q* and *S* points are detected. The minimum value of binary to the left and right is detected. This minimum value is taken as *Q* and *S* points. This results in binary string and can be taken as a signature, and can act as a hash value in phase 1. This signature is of *W* bytes. The next signature is considered to be part of the next *L* bytes. Figure 1 shows the function blocks of phase 1.

Phase 1 hash value is taken to generate the hash value for phase 2, which is of size m as shown in Figure 1. Similar operation is performed for phase 2. The array i is indexed by the hash value. Figure 1 also shows the working of phase 2. The counters (black locations in Figure 1) are incremented. The independent value of hash is calculated after calculating the final result for phase 1. Samples between points *Q* and *S* act as an independent hash value, i.e. m and it works till the last *R*-peak. Table 1 shows the number of discrete values. This sample acts as a hash value for phase 2, then binary addition is performed. Resultant value of binary addition acts as a memory address location for the array i.

If any counting unit crosses the fixed limit, the binary series is considered as a worm, i.e. QRS waveform. System setting is discussed as follows:

- $L \approx 288–432$: L is a notation and it depends on the start and end limits.
- $W \approx 12–28$: W is a notation of the position of points *Q* and *S*.
- $l = 8$: is in bits to check the limit in phase 1.
- $m \approx 12–21$: width of second phase of hashing.
- ≥ 12: addition of sample between points *Q* and *S*.
- k: limit of counting unit to reach the fixed value.
- HR: limit of counter value to be reached.

Based on the contents of the counter as shown in Table 2, a Gaussian curve is obtained, where irregularity lies on the end of the curve, as also discussed by Faezipour *et al.*[24].

For phase 3, the waveforms of different classes of heartbeat or groups of different classes of heartbeat are taken from the database of MIT–BIH to determine the qualitative features of the respective classes as discussed above, so that the system can be fine-tuned for any particular class. We tuned our system for classifying the waveform of ten arrhythmia classes: NORM, LBBB, RBBB, VPC, APC, NORM-RBBB, NORM-VPC, NORM-APC, LBBB-VPC and RBBB-APC. The paired features for a particular class found in phase 2 will now act as a hash function for classification of ECG waveform of a particular class in phase 3. As a number of data tables are used to identify the heartbeats, these data have been compressed (useful information between the paired features) to a single integer data value. The hash result would be the signature of phase 3. As a result, phase 3 has several hash functions which will now act as a signature for classification (paired features) obtained from phase 2 by the waveform of different classes of heartbeat. The MIT–BIH arrhythmia database can be applied as a signature and the counter used in parallel gives the result of the repeated hash value. We have selected the addition of the paired features as our primary hash function in phase 3. To obtain the integer hash value we round-off our results. Now the ECG waveform to classify the heartbeat is chosen. The profile curve for the respective hash function of phase 1 is plotted. If no anomalies lie on the tail(s), then we conclude that the ECG waveform belongs to normal class. On the other hand, if anomalies are found at the end of the curve, we conclude that the waveform has abnormal ECG beats and is forwarded to phase 3 for further processing. Then the profile curve for the respective hash function of phase 3 is plotted. Graph is plotted with memory location in x-axis and its content in y-axis. The graph will result in a inverted single bell-shaped curve or in a inverted double bell-shaped curve. Single curve means a single heartbeat class is present and double curve means that more than one class of heartbeat case is present. Therefore, if any of the first four hash functions of phase 3 result in a bell-shaped (single) profile curve, then

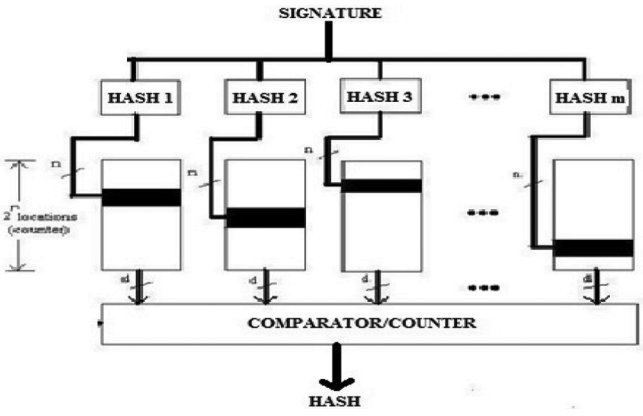

Figure 1. System design concept for phases 1 and 2.

Table 1. Result of QRS detection system when tested on MIT–BIH arrhythmia dataset 103 (10 s long) using the proposed algorithm

R-peak time	MIT–BIH *R*-peak sample index	*R*-peak index calculated	*R*–*R* interval	QRS duration (no. of samples)	Sum of *Q*–*S* array
0	21	0	0.736111	20	3584
0.736111	265	265	0.863889	18	3446
1.6	575	576	0.836111	19	3567
2.436111	876	877	0.844444	18	3411
3.280556	1180	1181	0.836111	18	3688
4.11667	1482	1482	0.897212	18	3516
5.013887	1795	1802	0.897111	20	3942
5.911111	2127	2128	0.897333	18	3379
6.788889	2444	2444	0.836111	20	3664
7.625	2744	2745	0.830556	19	3599
8.455556	3044	3044	0.841667	20	3767
9.297222	3347	3347	0.833333	20	3963
10.13056	3647	3647	0.883315	22	4161

Q and *S* are the points of the QRS complex in the ECG waveform.

Table 2. Storage unit content for profiling when tested on MIT–BIH arrhythmia database (15 min long each)

Datasets					
232		100		200	
No. of samples between *Q* and *S* points	Counts of sample	No. of samples between *Q* and *S* points	Counts of sample	No. of samples between *Q* and *S* points	Counts of sample
8	4	4	1	4	2
9	1	5	1	7	2
10	5	6	1	8	1
11	13	7	1	9	8
12	11	8	6	10	11
13	15	9	3	11	13
14	8	10	4	12	11
15	22	11	3	13	14
16	10	12	2	14	10
17	8	13	2	15	25
18	16	14	46	16	12
19	29	15	303	17	6
20	33	16	309	18	10
21	44	17	200	19	25
22	245	18	97	20	16
23	341	19	65	21	23
24	43	20	56	22	75
25	11	21	25	23	190
26	2	22	6	24	270
27	1	23	5	25	277
28	7	25	1	26	138
29	79	29	1	27	41
30	1			28	28
				29	205
				30	1

that waveform belongs to the heartbeat class LBBB or RBBB or VPC or APC respectively for which the particular hash function has resulted in a bell-shaped curve. Similarly, bell-shaped (double) profile curve for the next five hash functions of phase 3 results in the heartbeat class NORM-RBBB or NORM-VPC or NORM-APC or LBBB-VPC or RBBB-APC respectively.

We run our algorithm for data of first 15 min of MIT–BIH database. Irregular beats are generated at the end of the curve. Figure 2 shows how the contents of the memory elements get distributed for the datasets.

Normal distribution curve will result if we make a plot between count value and its content, i.e. count value plotted on the *x*-axis and number of times that count value is

Table 3. Results of area of the curve laying outside the Gaussian curve when tested on MIT–BIH arrhythmia database (15 min long each)

Dataset	MIT–BIH abnormal (%)	Total no. counts of the samples	$A = \sum_{x=a}^{b} y(x)$	Our abnormal (%)	Error (%)
103	0	1046	0	0	0
100	0.675	1138	7	0.615	0.06
200	19.34	1413	287	20.31	0.97
232	15.79	949	148	15.59	0.20
105	2.09	1305	12	0.919	1.17
112	0	1260	0	0	0
Avg.	6.3158			6.239	0.4

Table 4. Comparison of beat classification approaches on signal from MIT–BIH arrhythmia database

Reference	Algorithm	Se (%)	Sp (%)	No. of features
28	de Chazal *et al.*	77.7	98.8	15
29	Christov *et al.*	94.77	99.08	15
17	Iliev *et al.*	95.7	99.46	3
26	Besrour *et al.*	98.38	94.87	10–15
30	Haseena *et al.*		DER = 97.54	6
Proposed algorithm	–		DER = 99.57	1

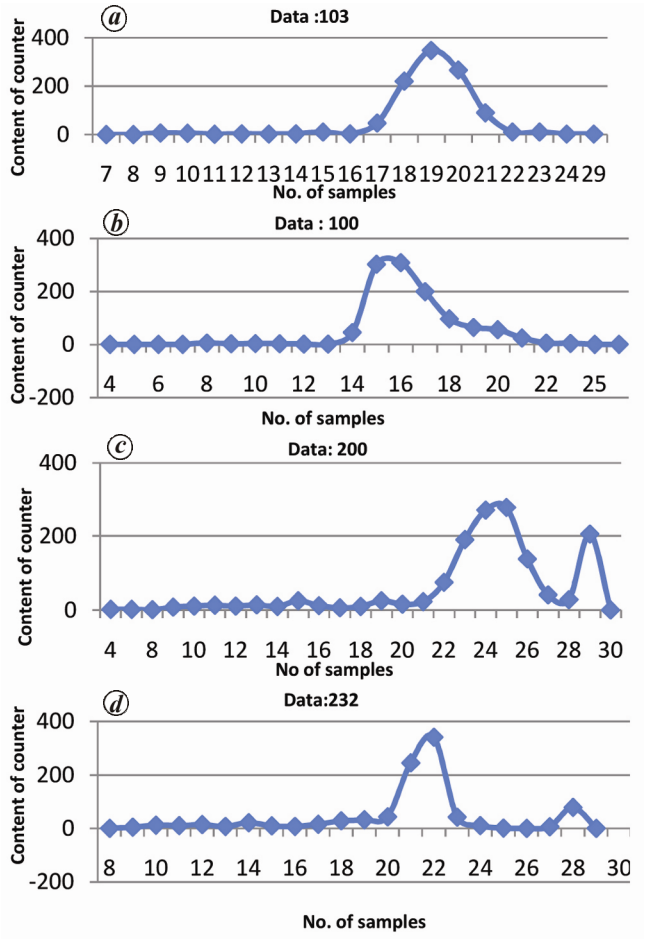

Figure 2. Distribution of counter contents for MIT–BIH arrhythmia database readings.

repeated is plotted on the *y*-axis (Figure 2). Any value which is not zero and lies after the Gaussian curve represents the irregular beats. To cross-check the results we calculated the area under the curve lying outside the Gaussian curve, and the results are shown in Table 3. Equation (1) gives the area under the curve for abnormal regions.

$$\text{Area} = \sum_{\text{Start point}}^{\text{End point}} \text{Content of counter lying outside Gaussian curve.} \tag{1}$$

Table 3 shows the performance of the proposed algorithm (regular and irregular beats). Point (16,309) in Figure 2 *b* shows that sample value 16 has the count value 309 in a ECG waveform of data number 100. So our result shows which beats are regular and which are irregular in the Gaussian curve.

Table 4 shows a comparison of the methods used for classification. Table 5 shows the five statistical indices used for comparison: sensitivity (Se), positive predictive value (PPV), specificity (Sp), negative predictive value (NPV) and total classification accuracy (TCA/DER). These are defined in eqs (2)–(6) below

$$\text{Se}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i}, \tag{2}$$

$$\text{PPV}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i}, \tag{3}$$

$$\text{Sp}_i = \frac{\text{TN}_i}{\text{TN}_i + \text{FP}_i}, \tag{4}$$

$$\text{NPV}_i = \frac{\text{TN}_i}{\text{TN}_i + \text{FN}_i}, \tag{5}$$

$$\text{DER} = \frac{\text{Correct detected beats}}{\text{Total beats}}. \tag{6}$$

Table 5. Performance evaluation of our cardiac profiling scheme for heartbeat classification on MIT–BIH arrhythmia database (15 min long each)

Dataset	MIT–BIH total beats	Detected total beats	$A = \sum_{x=a}^{b} y(x)$	MIT–BIH abnormal beats	Detected abnormal beats
103	1046	1046	0	0	0
100	1138	1138	7	8	7
200	1413	1413	287	273	287
232	949	949	148	150	148
105	1305	1305	12	27	12
112	1260	1260	0	0	0

Dataset	TP	FN	TN	FP	Se (%)	PPV (%)	Sp (%)	NPV (%)
103	1046	0	0	0	100	100	100	100
100	1130	0	7	1	100	99	87	100
200	1112	14	287	0	98.75	100	100	95.34
232	799	0	148	2	100	99.75	98.66	100
105	1278	0	12	15	100	98.83	44.44	100
112	1260	0	0	0	100	100	100	100
Avg.					99.79	99.59	88.35	99.22

Where TP, true positive; FN, false negative; FP, false positive; TN, true negative.

We proposed a method for detecting regular and irregular heartbeats in any individual. The method works on binary data information instead of any fixed value of features. Therefore any changes in amplitude and time do not affect the efficiency of the proposed method. We have set our system for classifying the waveform of ten arrhythmia classes. The output results in a Gaussian curve for a particular person, indicating the presence of irregular beats. The most important characteristic of the proposed method is that it is independent of gender (male or female) and age (young or old) and physical condition of any individual (walking, running, sleeping, etc.) because it uses binary data as information instead of any fixed value of the ECG waveform. The proposed method can be used as an alerting system as it detects irregular heartbeats in the ECG waveform. Such automated system will be useful in remote areas where a cardiologist may not be easily available and also for those who cannot afford expensive cardiac treatment.

1. Lilly, L. S., *Pathophysiology of Heart Disease*, Lippincott Williams & Wilkins, Philadelphia, USA, 2003, 3rd edn, pp. 57–90.
2. SNORT Network Intrusion Detection System; www.snort.org
3. MIT-BIH; http://www.physionet.org
4. Senhadji, L. *et al.*, Comparing wavelet transforms for recognizing cardiac patterns. *IEEE Eng. Med. Biol.*, 1995, **14**(2), 167–173.
5. Sahambi, J. S., Tandon, S. M. and Bhatt, R. K. P., Using wavelet transforms for ECG characterization: an on-line digital signal processing system. *IEEE Eng. Med. Biol.*, 1997, **16**(1), 77–83.
6. Hamilton, P. S. and Tompkins, W. J., Quantitative investigation of QRS detection rules using the MIT–BIH arrhythmia database. *IEEE Trans. Biomed. Eng.*, 1986, **33**(12), 1157–1165.
7. Thakor, N. V., Webster, J. G. and Tompkins, W. J., Estimation of QRS complex power spectra for design of a QRS filter. *IEEE Trans. Biomed. Eng.*, 1986, **31**(11), 702–706.
8. Li, C., Zheng, C. and Tai, C., Detection of ECG characteristic points using wavelet transforms. *IEEE Trans. Biomed. Eng.*, 1995, **42**(1), 21–28.
9. Castro, K. D. and Geva, A. B., ECG feature extraction using optimal mother wavelet. In IEEE EMBE International Conference, 2000, pp. 346–350.
10. Romero, I., Serrano, L. and Ayesta, ECG frequency domain features extraction: a new characteristic for arrhythmias classification. Conf. IEEE Eng. Med. Boil. Soc., 2001, p. 2.
11. Kadambe, S., Murray, R. and Boudreaux-Bartels, G. F., Wavelet transform based QRS complex detector. *IEEE Trans. Biomed. Eng.*, 1999, **46**(7), 838–848.
12. Romero, L., Addison, P. S. and Grubb, N., *R*-wave detection using continuous wavelet modulus maxima. *IEEE Proc. Comp. Cardiol.*, 2003, **30**, 565–568.
13. Tkacz, E. J. and Komorowski, D., An improved statistical approach to the QRS detection problem using matched filter facilities. *Biomed. Technik/Biomed. Eng.*, 1992, **37**(1), 99–109.
14. Pan, J. and Tompkins, W. J., A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.*, 1885, **32**(3), 230–236.
15. Dotsinsky, I. A. and Stoyanov, T. V., Ventricular beat detection in single channel electrocardiograms. *BioMed. Eng. Online,* 2004, **3**, 3.
16. Zhang, F., Tan, J. and Lian, Y., An effective QRS detection algorithm for wearable ECG in body area network. In Proceedings of IEEE Biomedical Circuits System Conference, 2007, pp. 195–199.
17. Iliev, I., Krasteva, V. and Tabakov, S., Real-time detection of pathological cardiac events in the electrocardiogram. *Physiol. Meas.*, 2007, **28**(3), 259–276.
18. Zhou, H.-Y. and Hou, K.-M., Embedded real-time QRS detection algorithm for pervasive cardiac care system. In Proceedings of 9th IEEE International Conference on Signal Process, 2008, pp. 2150–2153.
19. Cvikl, M. and Zemva, A., FPGA-oriented HW/SW implementation of ECG beat detection and classification algorithm. *Digital Signal Process*, 2010, **20**(1), 238–248.
20. Ning, X. and Selesnick, I. W., ECG enhancement and QRS detection based on sparse derivatives. *Biomed. Signal Process. Control.*, 2013, **8**, 713–723.
21. Jain, S., Kumar, P. and Subashini, M. M., LABVIEW based expert system for detection of heart abnormalities. In International Conference on Advances in Electrical Engineering, 2014, pp. 1–5.

22. Yazdani, S. and Vesin, J.-M., Adaptive mathematical morphology for QRS fiducial points detection in the ECG. *Comput. Cardiol.*, 2014, **38**, 725–728.
23. Nallathambi, G. and Principe, J. C., Integrate and fire pulse train automation for QRS detection. *IEEE Trans. Biomed. Eng.*, 2014; **61**(2), 317–326.
24. Faezipour, M., Nourani, M. and Panigrahy, R., A real-time worm outbreak detection system using shared counters. In Proceedings of the 15th Annual IEEE Symposium High Performing Interconnects, Dallas, 2007, pp. 65–72.
25. Song, M. H., Lee, J., Cho, S. P., Lee, K. J. and Yoo, S. K., Support vector machine based arrhythmia classification using reduced features. *Int. J. Control, Autom. Syst*., 2005, **3**(4), 571–579.
26. Besrour, R., Lachiri, Z. and Ellouze, N., ECG beat classifier using support vector machine. In Proceedings of the Third IEEE International Conference on Information and Communication Technology: From Theory Application, 2008, pp. 1–5.
27. Hu, Y. H., Palreddy, S. and Tompkins, W. J., A patient adaptive ECG beat classifier using a mixture of experts approach. *IEEE Trans. Biomed. Eng*., 1997, **44**(9), 891–900.
28. de Chazal, P., O'Dwyer, M. and Reilly, R. B., Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Trans. Biomed. Eng*., 2004, **51**(7), 1196–1206.
29. Christov, I., Gómez-Herrero, G., Krasteva, V., Jekova, I., Gotchev, A. and Egiazarian, K., Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification. *Med. Eng. Phys*., 2006, **28**(9), 876–887.
30. Haseena, H. H., Mathew, A. T. and Paul, J. K., Fuzzy clustered probabilistic and multi layered feed forward neural networks for electrocardiogram arrhythmia classification. *J. Med. Syst*., 2009, **29**(4), 1–10.
31. Chandrakar, C. and Sharma, M., A real time approach for ECG signal denoising and smoothing using adaptive window technique. In IEEE 9th International Conference on Industrial and Information Systems, Indian Institute of Information Technology, Gawlior, 2014, pp. 1–6.

Received 14 March 2016; accepted 12 November 2016

doi: 10.18520/cs/v112/1915-1920

A new formulation for determination of the competition coefficient in multispecies interaction for Lotka–Volterra type competition models

Anshuman Swain*,† and Saswata Chatterjee†

Undergraduate Department, Indian Institute of Science, Bengaluru 560 012, India

Determination of competition coefficients constitutes a vital part in the competition-based Lotka–Volterra-type population dynamics models. Various models have been proposed for the same, some of which were instinctive formulations, while some others were derived from dynamical and equilibrium relations pertaining to population dynamics. In this work, a new instinctive formulation to determine the competition coefficient has been proposed based on various parameters that determine the intensity of interspecific competition like the availability of resources, relative importance of a particular resource for a species, energy expenditure per resource utilization, etc.

Keywords: Competition, interspecies interaction, niche overlap, resource utilization.

THE study of interspecific competition amongst species has been one of the most prevalent concerns in the field of ecology, since the conception of competition. Most of the early evidence was forged on the basis of observational cues and limited laboratory experimentation (that too under constrained circumstances), rather than on-the-field studies[1]. This led to the questioning of the very principle of competition being a major force in population dynamics, especially in the 1960s–80s, resulting in a surge of experiments to prove or disprove the idea.

Most of the attacks were invalidated, but not without the amendment of the original theory of competition[1]. Schoener[1] reviewed about 164 experiments and discussed the results, which showed that interspecific competition occurred consistently in different habitat types and also at varied trophic levels as predicted by the theory. However, these experiments were designed to minimize the effect of predation, and thus did not account for the predatory effect. Later, Sih *et al.*[2] and Gurevitch *et al.*[3] analysed numerous experiments which had taken both predation and competition into account. This analysis demonstrated predation as a stronger effect, but nevertheless confirmed competition as a powerful factor in the interactions of multiple species[2]. Even today, numerous experiments[4,5] and simulations[6] demonstrate the importance of interspecific and intraspecific competition in modelling the dynamics of populations and their evolution.

The equations for interspecific competition, as suggested by Volterra[7] and later expounded by Lotka and Gause[8], which form the basis of our discussion, are of the general linear form (for n species)

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = r_i N_i \left[1 - \frac{\sum_{j=1}^{n} \alpha_{ij} N_j}{K_i} \right], \tag{1}$$

where n is the total number of interacting species, N_i the population size of species i, K_i the carrying capacity of species i, α_{ij} the competition coefficient of species i due to species j and r_i is the intrinsic growth rate of species i. Here, the competition coefficient is a key phenomenological measure of the interspecific interaction and serves an important part in the modelling of actual dynamics. (Note that nonlinear forms of these models also exist.)

*For correspondence. (e-mail: anshuman@ug.iisc.in)
†Equal contributors.