the empirical cumulative distribution functions (ECDFs) of temperature and galvanic skin response data were estimated separately for each intensity category, and the three breakpoints $(Be_{i,i=1:3})$ that divided the data into four equiprobable partitions $(Pe_{j,j=1:4})$ were derived. Separately the breakpoints $(Bd_{i,i=1:3})$ which divided the same set of data in four partitions $(Pd_{j,j=1:4})$ minimizing the mean square distortion of the quantization [155] were also calculated. The final partition breakpoints $(Bf_{i,i=1:3})$ were calculated averaging the corresponding pairs of breakpoints $Bf_i = \frac{Be_i + Bd_i}{2}$ and used to divide the data into four contiguous, non-overlapping ranges of values. Final partition ranges $(Pf_{j,j=1:4})$ are sorted in ascending order such that the first range (Pf_1) represents partition of data with the smallest values, and the last range (Pf_4) represents data with the highest values. The vocabulary of terms was built by allowing all the possible combinations between partitions ranges of temperature, galvanic skin response data and binary values of steps that share the same *IC*.

For the sleeping category, for example, the 32 terms of the vocabulary describing the sleeping intensity category can be represented by:

$$\begin{array}{lllll} \mathrm{t}_1^{\mathrm{S}}: & [\mathrm{S}, & \mathrm{Pf}_1^{\mathrm{ST}}, & \mathrm{Pf}_1^{\mathrm{GSR}}, & \mathrm{Steps}_{\mathrm{No}}] \\ \mathrm{t}_2^{\mathrm{S}}: & [\mathrm{S}, & \mathrm{Pf}_1^{\mathrm{ST}}, & \mathrm{Pf}_1^{\mathrm{GSR}}, & \mathrm{Steps}_{\mathrm{Yes}}] \\ & & \vdots & & \\ \mathrm{t}_{32}^{\mathrm{S}}: & [\mathrm{S}, & \mathrm{Pf}_4^{\mathrm{ST}}, & \mathrm{Pf}_4^{\mathrm{GSR}}, & \mathrm{Steps}_{\mathrm{No}}] \end{array}$$

The sum of terms across different activity levels is the total number of words. In particular, a total of 128 terms were initially created. As a last step, in order to remove nonsense words and increase the frequency of *VL* and *MV* words to obtain sparse topics, we pruned the vocabulary adding wildcard characters [138]. We used a wildcard character to replace the symbol related to the steps performed during sleeping *IC* (i.e. 16 words removed). Considering the neutral wildcard symbols the original 32 terms representing the sleeping category can be represented by 16 terms as:

$$\begin{array}{llll} \mathrm{t}_1^{\mathrm{S}}: & [\mathrm{S}, & \mathrm{Pf}_1^{\mathrm{ST}}, & \mathrm{Pf}_1^{\mathrm{GSR}}] \\ \mathrm{t}_2^{\mathrm{S}}: & [\mathrm{S}, & \mathrm{Pf}_1^{\mathrm{ST}}, & \mathrm{Pf}_2^{\mathrm{GSR}}] \\ & & \vdots & \\ \mathrm{t}_{16}^{\mathrm{S}}: & [\mathrm{S}, & \mathrm{Pf}_4^{\mathrm{ST}}, & \mathrm{Pf}_4^{\mathrm{GSR}}] \end{array}$$

Two wildcard characters replaced the temperature and galvanic skin response symbols during *VL* and *MV* intensities (i.e. 60 words removed) to know if the subject was in these two *IC* because of moving. In view of sparsity we also weighted the informativeness of the remaining words in the vocabulary based on their inverse document frequency (IDF) score. Those terms that have a high IDF are considered more informative, because they rarely occur in the collection. In particular, we removed the words occurring in at least 90% of the documents since, occurring so frequently, they are more likely to obscure than facilitate a meaningful decomposition of the collection of documents. The term frequency (TF), usually used in combination with IDF to form the TF-IDF score [141], was not considered since it penalizes words that rarely appear within a document such as words of light or moderate to vigorous intensity, these words are important in the identification of sleep modalities