shown in FIG. 6, according to the rule preset on the deduplication server 103, a point p_{j1} and a window $W_{j1}[p_{j1}\text{-}169, p_{j1}]$ corresponding to the point p_{j1} are determined for the potential dividing point k_j. A manner of determining whether at least a part of data in the window $W_{j1}[p_{j1}\text{-}169, p_{j1}]$ meets the preset condition C_1 is the same as the manner of determining whether at least a part of data in the window $W_{i1}[p_{i1}\text{-}169, p_{i1}]$ meets the preset condition C_1. Therefore, as shown in FIG. 17, W_{j1} represents the window $W_{j1}[p_{j1}\text{-}169, p_{j1}]$, and 5 bytes are selected to determine whether the at least a part of data in $W_{j1}[p_{j1}\text{-}169, p_{j1}]$ meets the preset condition C_1. In FIG. 17, each of bytes "■" whose sequence numbers are 169, 127, 85, 43, and 1 represents 1 selected byte, and there are 42 bytes between two adjacent selected bytes. The bytes "■" whose sequence numbers are 169, 127, 85, 43, and 1 are regarded as 40 sequential bits, which are represented as a_1', a_2', a_3', a_4', ..., and a_{40}' separately. For any a_t' of a_1', a_2', a_3', a_4', ..., and a_{40}', when $a_t'=0$, $V_{at}'=-1$, and when $a_t'=1$, $V_{at}'=1$. V_{a1}', V_{a2}', V_{a3}', V_{a4}', ..., and V_{a40}' are generated according to a correspondence between a_t' and V_{at}'. The manner of determining whether at least a part of data in the window $W_{j1}[p_{j1}\text{-}169, p_{j1}]$ meets the preset condition C_1 is the same as the manner of determining whether at least a part of data in the window $W_{i1}[p_{i1}\text{-}169, p_{i1}]$ meets the preset condition C_1, and therefore, the same random numbers are used: h_1, h_2, h_3, h_4, ..., and h_{40}. $S_a'=V_{a1}'*h_1+V_{a2}'*h_2+V_{a3}'*h_3+V_{a4}'*h_4+...+V_{a40}'*h_{40}$. Because h_1, h_2, h_3, h_4, ..., and h_{40} follow normal distribution, S_a' also follows normal distribution. When S_a' is a positive number, the at least a part of data in $W_{j1}[p_{j1}\text{-}169, p_{j1}]$ meets the preset condition C_1; when S_a' is a negative number or 0, the at least a part of data in $W_{j1}[p_{j1}\text{-}169, p_{j1}]$ does not meet the preset condition C_1. A probability that S_a' is a positive number is 1/2.

[0137] The manner of determining whether at least a part of data in $W_{i2}[p_{i2}\text{-}169, p_{i2}]$ meets the preset condition C_2 is the same as a manner of determining whether at least a part of data in $W_{j2}[p_{j2}\text{-}169, p_{j2}]$ meets the preset condition C_2. Therefore, as shown in FIG. 17, "▩" represents 1 byte selected when it is determined whether the at least a part of data in the window $W_{j2}[p_{j2}\text{-}169, p_{j2}]$ meets the preset condition C_2, and there are 42 bytes between two adjacent selected bytes. In FIG. 17, selected bytes are represented as sequence numbers 170, 128, 86, 44, and