

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] The foregoing and other features and advantages of the invention will be apparent from the following, more particular description of various embodiments, as illustrated in the accompanying drawings wherein like reference numbers generally indicate identical, functionally similar, and/or structurally similar elements. The first digits in the reference number indicate the drawing in which an element first appears. Unless otherwise indicated, the accompanying drawing figures are not to scale.

[0007] FIG. 1 depicts a sample histogram of eleven units;

[0008] FIG. 2 depicts a sample time-series, with the measurement value on the y-axis and a time of measurement on the x-axis;

[0009] FIG. 3A-3D depict histograms using color density;

[0010] FIG. 4 depicts side-by-side vertical histograms representing measurements over time;

[0011] FIG. 5 depicts example fixed bin sizes;

[0012] FIG. 6 depicts an example of multiple bin aggregation;

[0013] FIG. 7 depicts an example flowchart describing processing performed in an illustrative embodiment of the present invention; and

[0014] FIG. 8 depicts an example computer processing system that may be used in implementing an embodiment of the present invention.

DETAILED DESCRIPTION

[0015] Illustrative embodiments are discussed in detail below. While specific embodiments are discussed, it should be understood that this is done for illustration purposes only. In describing and illustrating the embodiments, specific terminology is employed for the sake of clarity. However, the embodiments are not intended to be limited to the specific terminology so selected. A person skilled in the relevant art will recognize that other components and configurations may be used without departing from the spirit and scope of the embodiments. It is to be understood that each specific element includes all technical equivalents that operate in a similar manner to accomplish a similar purpose. The examples and embodiments described herein are non-limiting examples.

[0016] Furthermore, the embodiments detailed below may be combined into a new embodiment and/or various features of the embodiments described below may be combined to form a new embodiment.

[0017] All publications cited herein are hereby incorporated by reference in their entirety. As used herein, the term "a" refers to one or more. The terms "including," "for example," "such as," "e.g.," "may be" and the like, are meant to include, but not be limited to, the listed examples.

[0018] FIG. 1 depicts a sample histogram of eleven units or bins. A histogram is a dense representation of data samples where samples themselves may be grouped together (losing accuracy). For example, given a set of numbers 0.5, 0.7, 2.0, 2.9, 7 and 9.2, one may divide the space from 0 to 10 into units of 1. The histogram of FIG. 1 depicts 11 "bins" of size 1: [0,1), [1,2), [2,3), [3,4), [4,5), [5,6), [6,7), [7,8), [8,9), and [9,10). The number and size of the bins may change. There are two numbers in each of [0,1) and [2,3), and one number in each of [7,8) and [9,10) as shown in FIG. 1.

[0019] FIG. 2 depicts a sample time-series, with a measurement on the y-axis and a time the measurement was taken on the x-axis. A time series is a set of measurements over time. In this example, it is measured how long a specific action took (e.g. loading a web page) every minute for 30 minutes. A common visual representation (line graph) would be to place the measurement on the y-axis and the time on the x-axis as shown in FIG. 2.

(e.g., 16 bits) representation. The fixed-bit width may, for example, allow for more efficient key lookup. The bin size may be fixed (e.g., two significant digits base 10) regardless of the data values. This creates a uniform bin scheme that may be used, for example, across different data streams and keeps the same key space. In one embodiment, a fixed key space may be provided. The fixed key space is above and beyond fixed bins and may be a limitation to the number of total bins. The fixed key space may represent buckets between 1.0×10^{-127} and 9.9×10^{128} , 0, and between -1.0×10^{-127} and -9.9×10^{128} and thereby limits the keys represent into 16 bits.

[0033] The fixed key space allows for aggregating, for example, the 720 one-minute histograms into a single 12-hour histogram both (1) efficiently and (2) in a fashion that tracks error consistently. Histogram may be aggregated more efficiently because the keyspaces between the histograms are identical. For example, two histograms may be simply be added together (e.g., bin by bin). If the keyspaces were different, mathematical conversions would be required and the effects on error would be even more complicated. Additionally, the uniform binning across various data streams also provides for highly compressible histograms. For example, restricted use of key space leads to better compression.

[0034] Combining histograms with different bin selections may be an issue because different bin selections that overlap but are not equal have different error introduced when samples are placed within them. Previously, when a new aggregate bin was created, a technique to re-bin the samples using some selected distribution assumption must be performed. The technique may not be possible if the bins are either really large (e.g., not useful) or they are small but samples can be entirely misplaced (e.g., the possible error would exceed the bin itself).

[0035] In one embodiment, universal compatibility regardless of measurement magnitude while maintaining a satisfactory and useful level of precision may be accomplished through the following. First, the number zero may be treated as an exact bin (e.g., all measurements that equal zero). A bin zero may be defined as containing all measurements that share equality with the value zero. Non-zero numbers may be represented in bins that are bounded numbers with two significant digits in base ten. This has characteristics of representing numbers of wide magnitude space (i.e., 10^{-128} to 10^{127}) logarithmically, while maintaining a linear granularity within each bin set aiding human understanding and visualization (e.g., a large set of bins such as, [1.0, 1.1), [1.1, 1.2), . . . [9.9, 10) collectively times 10^n). Bins may be log-linear bins using 2 significant digits base 10, and, therefore, the maximal error is 5% and the average error is approximately 0.65%. As the base or significant digits changes, the error would change as well.

[0036] Bins may be from one point on a discrete number line to the next, with a closed interval on the point closest to zero and an open interval on the number furthest from zero (e.g. [1.7, 1.8) and [-3400, -3500) and [8.2.times.10¹⁴, 8.3.times.10¹⁴)). When two bins touch (e.g., two adjacent bins), for example, [1.0, 1.1) and [1.1, 1.2) the bins cannot both represent 1.1 (e.g., no overlap). If both bins were closed, both would represent the point where they touch. If both bins were open, neither would represent the point where they touch. Accordingly, one bin must be open and one bin must be closed. A convention is needed to determine which bin the value of 1.1 would go into. So, closed-interval includes the point, open does not. Accordingly, in one embodiment, the point closest to zero is closed. In Mathsquare, for example, brackets [] indicate closed (including) and parenthesis () indicate open (excluding).

[0037] FIG. 5 depicts a bin selection using two significant digits in base 10 and illustrates how the bottom 10% of each exponent range is actually represented by the adjacent smaller exponent. Additionally, the last exponent bin range in the figure depicts 0 as a special value wherein smaller measurements cannot be accounted for. This is a log-linear bin selection from 1.9×10^{-127} to 9.9×10^{128} .

[0038] Given discrete representation, selected histograms may be aggregated without arbitrarily expanding the bin space, which may provide an ability to aggregate (e.g., via addition and subtraction) while controlling bin space. This may have the distinct real-world advantage of a highly controlled bin space for a physical system (e.g., a system having prescribed constraints). FIG. 6 depicts an example of multiple bin aggregation. For example, histograms in (a) added with histograms in (b) equals the histograms in (c).

[0039] Techniques shown below calculate and store (e.g., time) aggregated histogram data to decrease required information access on retrieval.

[0040] In one embodiment, the display may communicate the number of histograms desired and the storage system may dynamically aggregate requested data into lengthier periods of time. For example, displaying a two-year histogram time series where all histograms are stored by minute would require 1,051,200 histograms returned to the visualization system. Instead, for example, the display may request 1,000 histograms, which may result in the storage systems combining 720 one-minute histograms inclusively between [00:00 and 11:59] into one 12-hour histogram and the 720 one-minute histograms inclusively between [12:00 and 23:59] into one 12-hour histogram for each day of the two years resulting in 1,460 histograms to the visualization system instead of 1,051,200 with conventional techniques. This new representation presents a waste of 46% as opposed to a waste of 105000% with conventional techniques.

[0041] In one embodiment, the display may understand horizontal resolution and determine how many pixels are available for rendering vertical histograms. This "horizontal viewport size" may be the number of histograms requested (e.g., a number equal to or less than the number of horizontal pixels). The histogram delivery technique may attempt to aggregate the number of bins to a level where the number of histograms returned for the requested time range roughly approximates but is less than the "horizontal viewport size" (e.g., within a factor of 4 to be roughly approximate).

[0042] In one embodiment, to determine the number of bins requested, available visual space may be pulled or requested from the display itself through various techniques. In a web browser, for example, the number of horizontal and vertical pixels available in the viewport may be accessible by calling a Javascript routine.

[0043] In one embodiment, the number of histograms may be less than the number of horizontal pixels, as any number greater may be lost during visualization. The number of bins desired may be less than the number of vertical pixels, as any number greater may be lost during visualization.

[0044] In one embodiment, the number of histograms is roughly half the number of horizontal pixels. For example, with 1000 pixels of horizontal space, roughly 500 histograms may be created. And, with 600 pixels of vertical space, 300 bins may be used.

[0045] In one embodiment, fine granularity histograms may be automatically aggregated into successively wider granularity histograms to accelerate visualization of data over arbitrary periods of time (e.g., storing 1 minute, 5 minute, 1 hour, and 12 hour aggregations). For example, each time a one-minute histogram is stored, the containing five-minute histogram is recalculated and stored, causing the containing one-hour histogram to be recalculated and stored, causing the containing 12-hour histogram to be recalculated and stored, etc. Accordingly, histogram data is pre-calculated as data is received. The largest feasible pre-calculated histogram accumulates is selected above. A user may provide the largest feasible pre-calculated histogram based on how often the user believes they would be used. For example, a user may provide the largest feasible pre-calculated histogram based on how long they believe they will retain data and how frequently long time periods will be visualized.

[0046] Bin-Realm Efficiency

[0047] Another advantage of the techniques described herein is an ability to represent histograms in a smaller information footprint based on the visualization parameters.

[0048] Human beings often interpret logarithmically scaled data incorrectly and tend to operate more effectively with linearly scaled data. Given a wide magnitude space and a linear representation, there may be too much data to display (e.g., 45,391 bins. When visualizing data, the resolution at which data may be usefully displayed may be determined. For example, on displays available today, humans cannot distinguish in excess of 1,000 histogram bins without external apparatus. The technique described below provides a method of dynamic rebinning that may optimize the histogram for visual delivery in linearly scaled representation where bins otherwise not visible on-display may be combined into a single bin.

[0049] In one embodiment, specific user-supplied minimum and maximum visual limits may be provided. For example the number of pixels and size of a display may impact the visual limits. The greater of the absolute value of the minimum and maximum visual limits is taken as the focal maximum. In one example, there may be 90 bins like [1.0,1.1), [1.1,1.2), . . . [9.9,10) each of which times 10 n. The logarithmic bin set (e.g., the 90 bins for a particular value of n) containing the focal maximum is preserved (e.g., unchanged,

(RAM), and a secondary memory 810, etc., may be a computer-readable medium that may be configured to store instructions configured to implement one or more embodiments and may comprise a random-access memory (RAM) that may include RAM devices, such as Dynamic RAM (DRAM) devices, flash memory devices, Static RAM (SRAM) devices, etc.

[0062] The secondary memory 810 may include, for example, (but is not limited to) a hard disk drive 812 and/or a removable storage drive 814, representing a floppy diskette drive, a magnetic tape drive, an optical disk drive, a compact disk drive CD-ROM, flash memory, etc. The removable storage drive 814 may, e.g., but is not limited to, read from and/or write to a removable storage unit 818 in a well-known manner. Removable storage unit 818, also called a program storage device or a computer program product, may represent, e.g., but is not limited to, a floppy disk, magnetic tape, optical disk, compact disk, etc. which may be read from and written to removable storage drive 814. As will be appreciated, the removable storage unit 818 may include a computer usable storage medium having stored therein computer software and/or data.

[0063] In alternative illustrative embodiments, secondary memory 810 may include other similar devices for allowing computer programs or other instructions to be loaded into computer system 800. Such devices may include, for example, a removable storage unit 822 and an interface 820. Examples of such may include a program cartridge and cartridge interface (such as, e.g., but not limited to, those found in video game devices), a removable memory chip (such as, e.g., but not limited to, an erasable programmable read only memory (EPROM), or programmable read only memory (PROM) and associated socket, and other removable storage units 822 and interfaces 820, which may allow software and data to be transferred from the removable storage unit 822 to computer system 800.

[0064] Computer 800 may also include an input device 803 which may include any mechanism or combination of mechanisms that may permit information to be input into computer system 800 from, e.g., a user or operator. Input device 803 may include logic configured to receive information for computer system 800 from, e.g. a user or operator. Examples of input device 803 may include, e.g., but not limited to, a mouse, pen-based pointing device, or other pointing device such as a digitizer, a touch sensitive display device, and/or a keyboard or other data entry device (none of which are labeled). Other input devices 803 may include, e.g., but not limited to, a biometric input device, a video source, an audio source, a microphone, a web cam, a video camera, and/or other camera.

[0065] Computer 800 may also include output devices 815 which may include any mechanism or combination of mechanisms that may output information from computer system 800. Output device 815 may include logic configured to output information from computer system 800. Embodiments of output device 815 may include, e.g., but not limited to, display 801, and display interface 802, including displays, printers, speakers, cathode ray tubes (CRTs), plasma displays, light-emitting diode (LED) displays, liquid crystal displays (LCDs), printers, vacuum fluorescent displays (VFDs), surface-conduction electron-emitter displays (SEDs), field emission displays (FEDs), etc. Computer 800 may include input/output (I/O) devices such as, e.g., (but not limited to) input device 803, communications interface 824, connection 828 and communications path 826, etc. These devices may include, e.g., but are not limited to, a network interface card, onboard network interface components, and/or modems.

[0066] Communications interface 824 may allow software and data to be transferred between computer system 800 and external devices or other computer systems. Computer system 800 may connect to other devices or computer systems via wired or wireless connections. Wireless connections may include, for example, WiFi, satellite, mobile connections using, for example, TCP/IP, 802.15.4, high rate WPAN, low rate WPAN, 6loWPAN, ISA100.11a, 802.11.1, WiFi, 3G, WiMAX, 4G and/or other communication protocols.

[0067] In this document, the terms "computer program medium" and "computer readable medium" may be used to generally refer to media such as, e.g., but not limited to, removable storage drive 814, a hard disk installed in hard disk drive 812, flash memories, removable discs, non-removable discs, etc. In addition, it should be noted that various electromagnetic radiation, such as wireless communication, electrical communication carried over an electrically conductive wire (e.g., but not limited to twisted pair, CAT5, etc.) or an optical medium (e.g., but not limited to, optical fiber) and the like may be encoded to carry computer-executable instructions and/or computer data that embodiments of the invention on e.g., a communication network. These computer program products may provide software to computer system 800. It should be noted that a computer-readable medium that comprises computer-executable instructions for execution in a

[Images](#)

[Add to Shopping Cart](#)

[View Shopping Cart](#)

[Hit List](#)

[Top](#)

[Help](#)

[Home](#)

[Boolean](#)

[Manual](#)

[Number](#)

[PTDLs](#)