



US 20080249834A1

(19) **United States**

(12) **Patent Application Publication**
Zigmond et al.

(10) **Pub. No.: US 2008/0249834 A1**

(43) **Pub. Date: Oct. 9, 2008**

(54) **ADJUSTING FOR UNCERTAINTY IN
ADVERTISEMENT IMPRESSION DATA**

(22) Filed: **May 1, 2007**

Related U.S. Application Data

(75) Inventors: **Daniel J. Zigmond**, Menlo Park,
CA (US); **Jason Bayer**, Mountain
View, CA (US); **Xiaohu Zhang**,
Mountain View, CA (US)

(60) Provisional application No. 60/909,893, filed on Apr.
3, 2007.

Publication Classification

(51) **Int. Cl.**
G06F 17/18 (2006.01)
G06Q 30/00 (2006.01)

(52) **U.S. Cl.** **705/10; 705/14**

(57) **ABSTRACT**

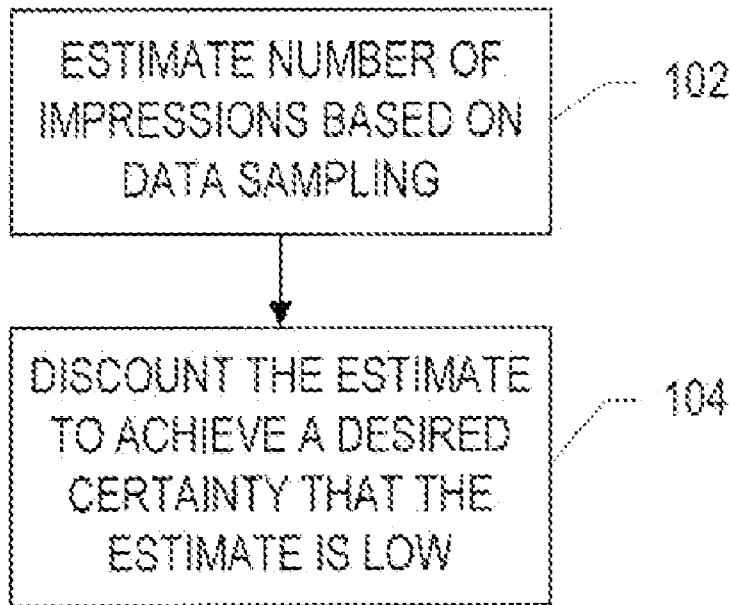
Correspondence Address:
FISH & RICHARDSON P.C.
PO BOX 1022
MINNEAPOLIS, MN 55440-1022 (US)

Methods, systems and apparatus, including computer program products, provide adjustments to advertisement impression estimates based on sampled data. According to one method, an estimated number of advertisement impressions is made based on the sampled data and the estimate is discounted to achieve a level of certainty that the resulting discounted estimate is less than the actual number of advertisement impressions. The amount of the discount can be determined based on a sampling error model of the sampling data.

(73) Assignee: **GOOGLE INC.**, Mountain View,
CA (US)

(21) Appl. No.: **11/743,078**

100 →



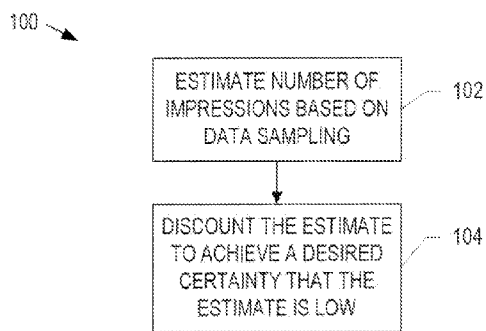


FIG. 1A

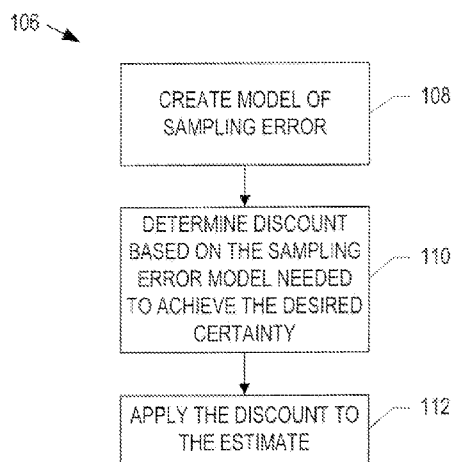


FIG. 1B

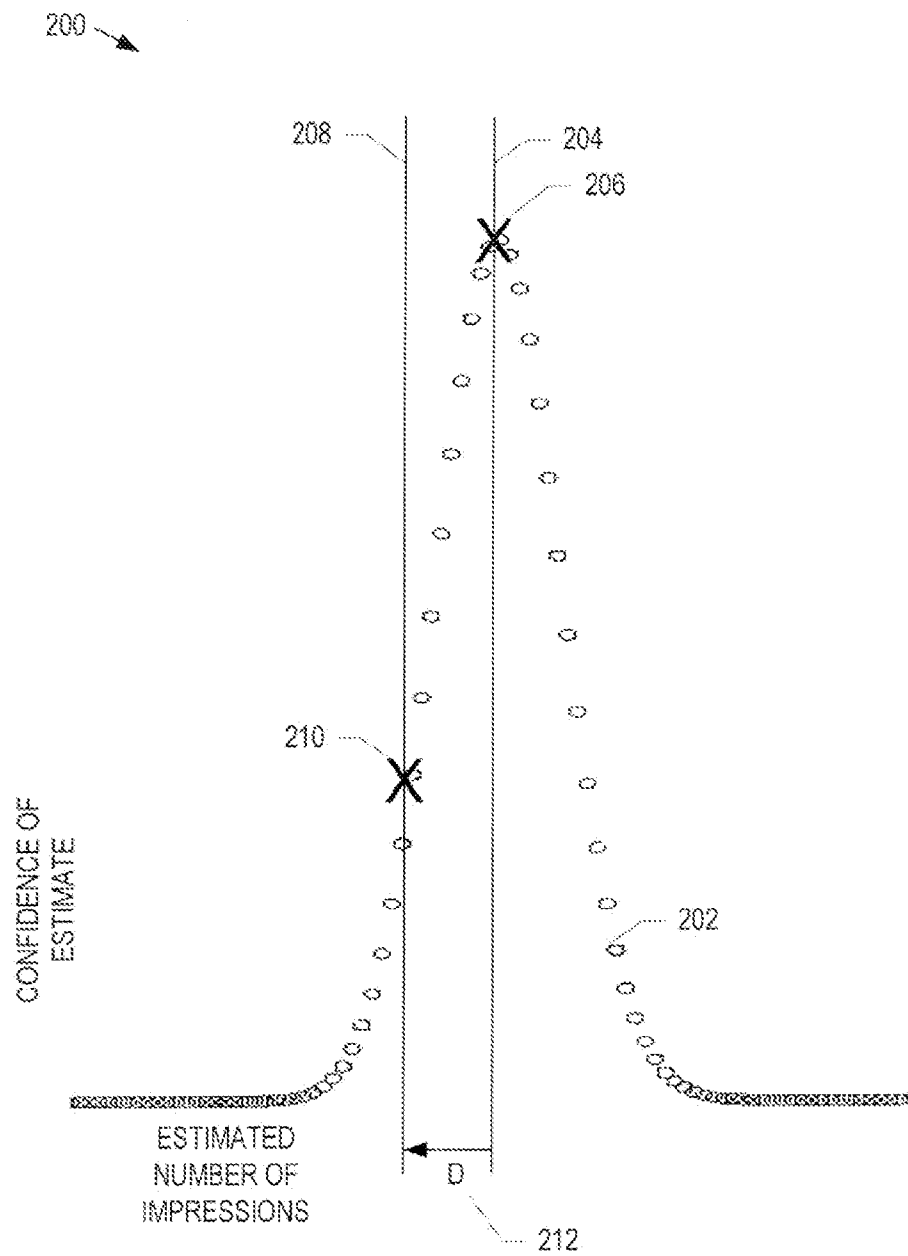


FIG. 2

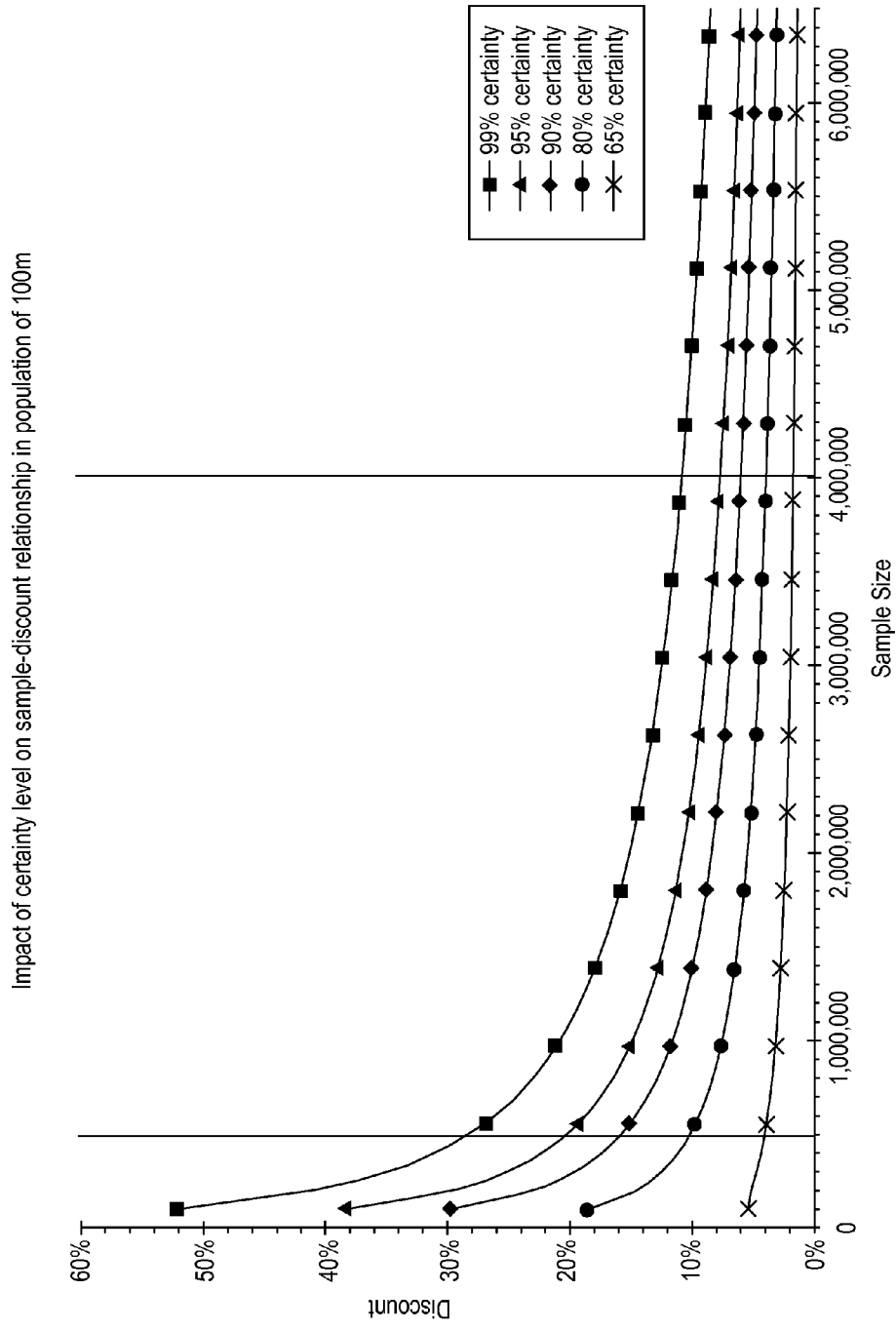


FIG. 3

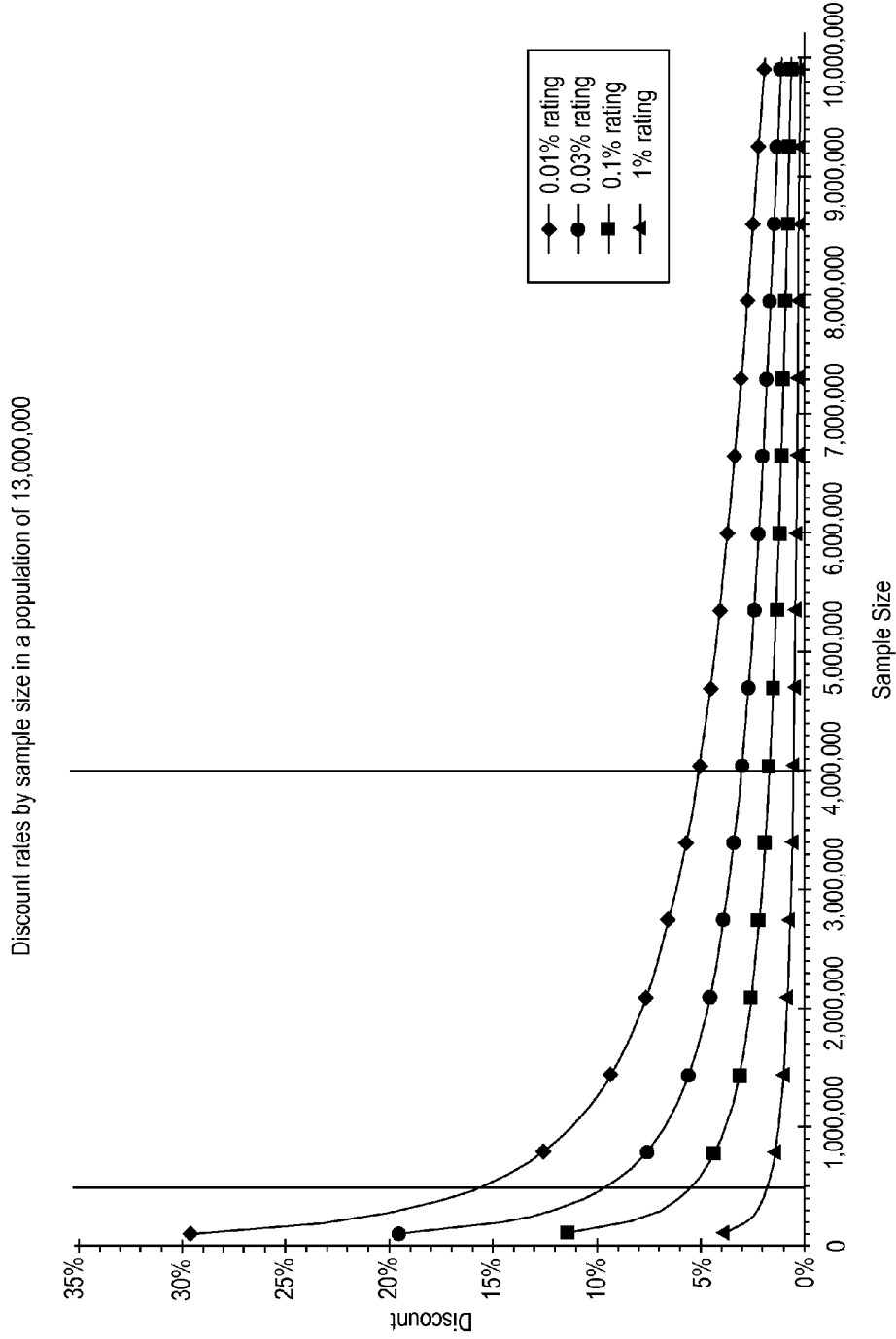


FIG. 4

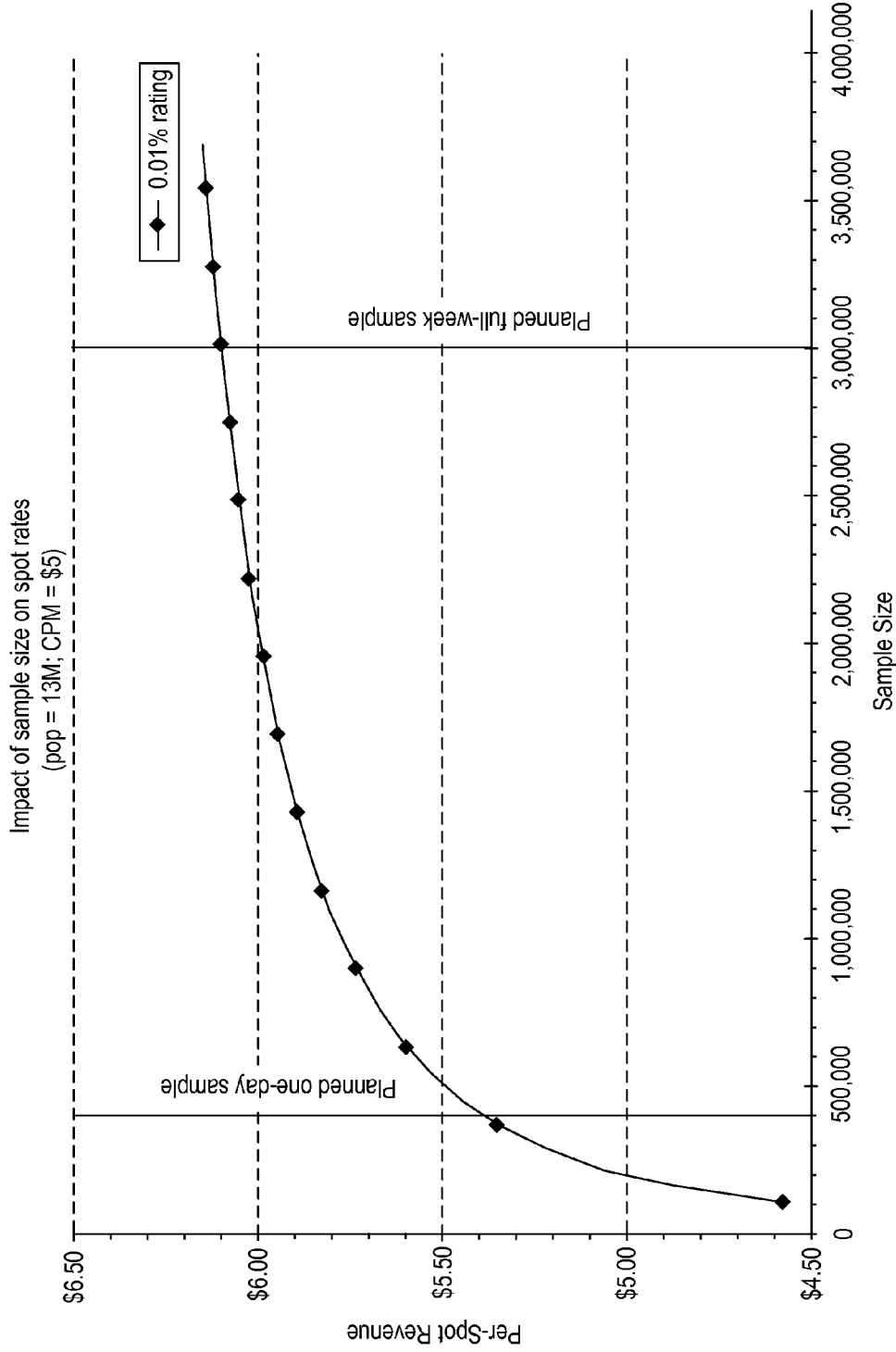


FIG. 5

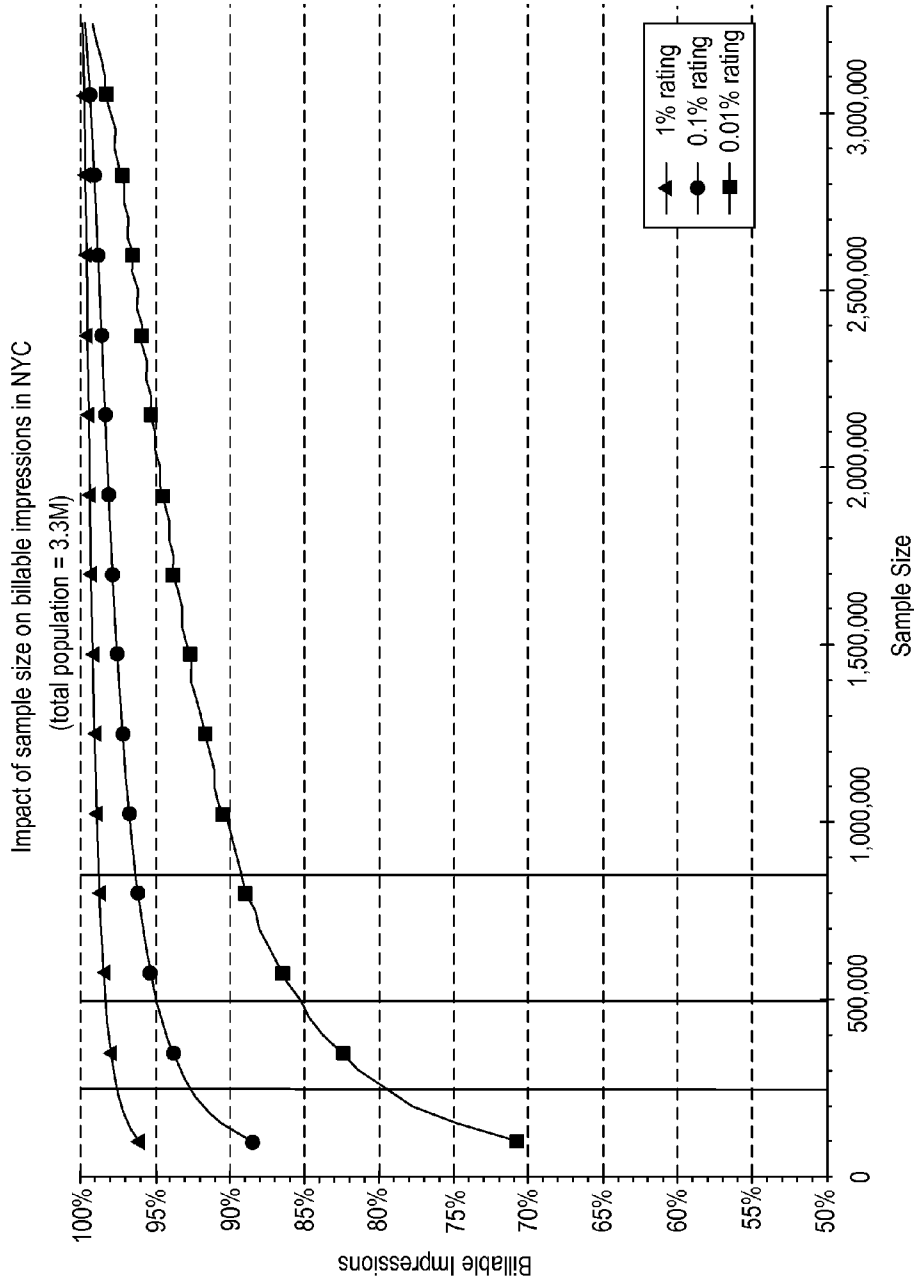


FIG. 6

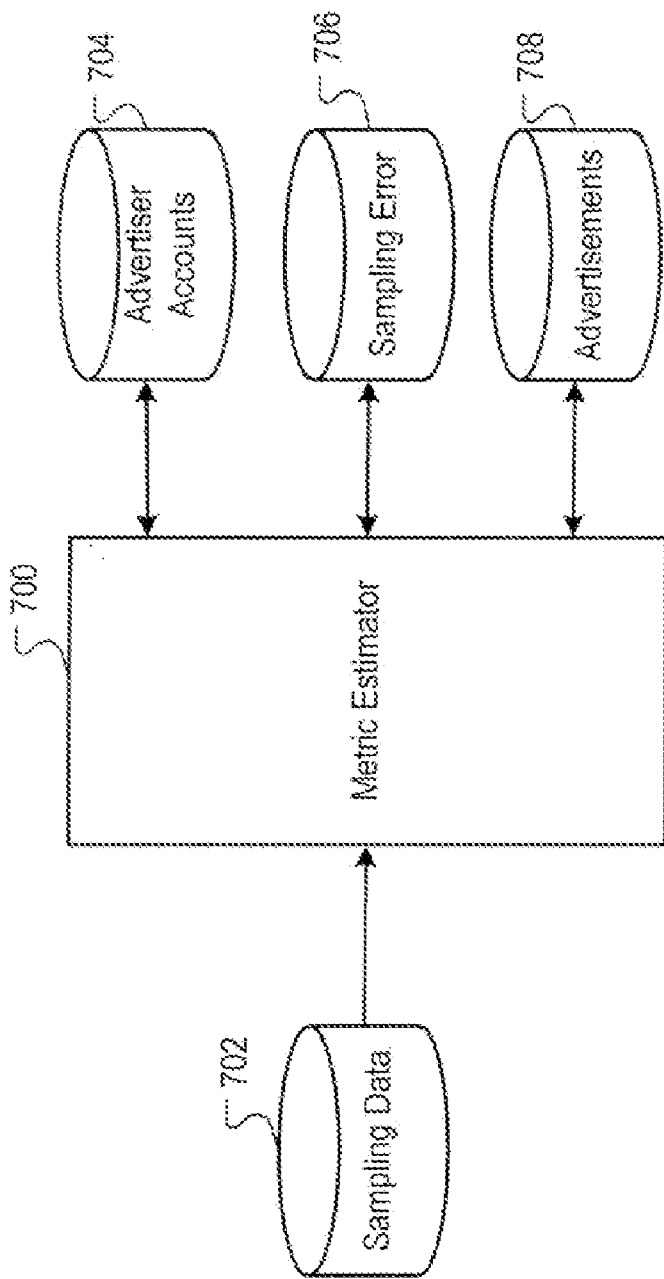


FIG. 7

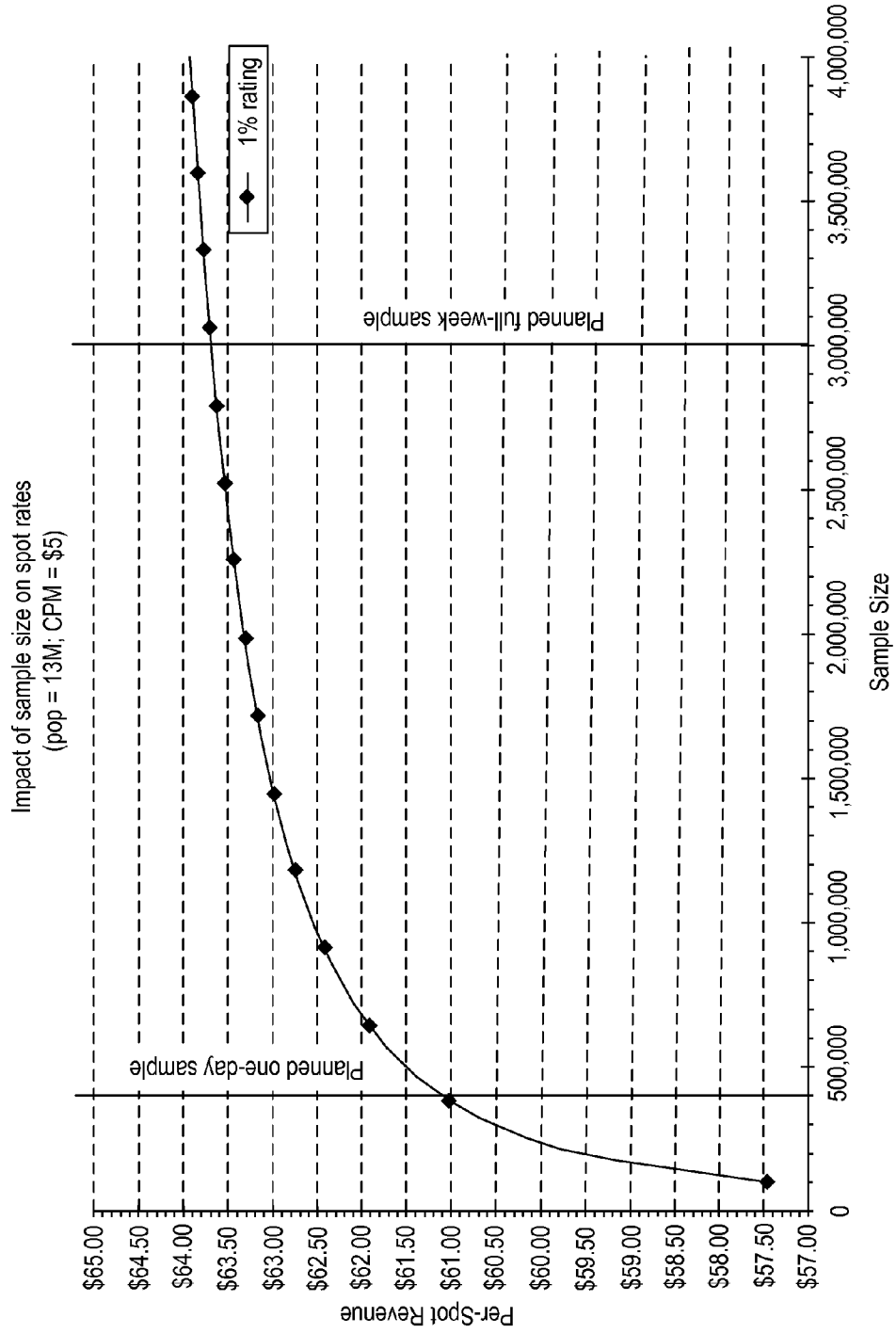


FIG. 8

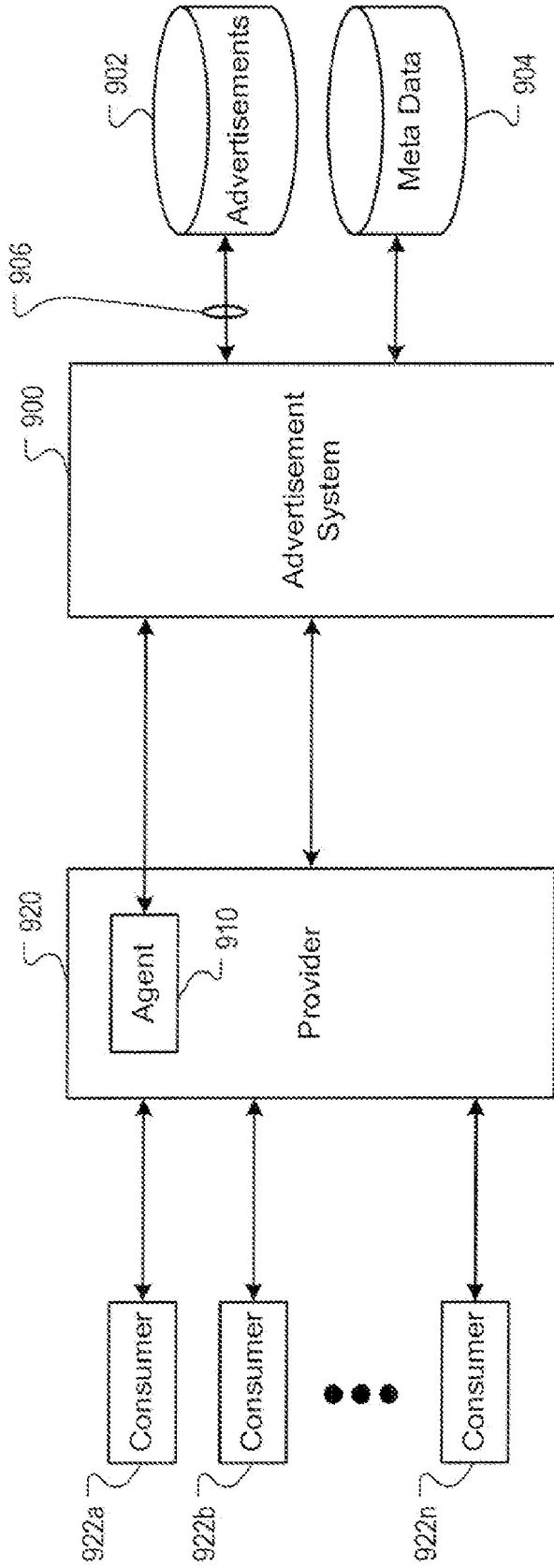


FIG. 9

**ADJUSTING FOR UNCERTAINTY IN
ADVERTISEMENT IMPRESSION DATA**

CROSS-REFERENCE

[0001] This application is claims priority to United States Provisional Patent Application Ser. No. 60/909,893, entitled "Television Advertising," filed on Apr. 3, 2007, the disclosure of which is incorporated herein by reference.

BACKGROUND

[0002] The present disclosure relates to advertising.

[0003] An advertiser, such as a business entity, can purchase airtime during a radio or television broadcast to air advertisements. Example radio and television advertisements include commercials that are aired during a program break. Television advertisements can additionally include overlays that are aired during a program, and text banners that are aired during a program.

[0004] The cost of the airtime purchased by the advertiser varies according to both the amount of time purchased and other parameters such as the audience size and audience composition expected to be listening or watching during the purchased airtime. The audience size and audience composition, for example, can be measured by a ratings system. Data for television ratings can, for example, be collected by viewer surveys in which viewers provide a diary of viewing habits; or by set meters that automatically collect viewing habit data and transmit the data over a wired or wireless connection, e.g., a phone line or cable line; or by digital video recorder service logs, for example. Such rating systems, however, may be inaccurate for niche programming, and typically provides only an estimate of the actual audience numbers and audience composition.

[0005] Based on the ratings estimate, airtime is offered to advertisers for a fee. Typically the advertiser must purchase the airtime well in advance of the time the airtime is broadcast. Additionally, the advertiser and/or the broadcaster may not realize the true value of the airtime purchased if the ratings estimate is inaccurate. Finally, the commercial that is aired during the airtime may not be relevant in the context of the television program or audience.

[0006] In order to make informed advertising strategy decisions, advertisers require information regarding the number of consumers to which a advertisement is delivered via a given mode of delivery (the ad's impression count). In many advertising contexts, measuring a commercial ad's impression count is difficult due to the nature of the advertising medium. In broadcast mediums, measurements are often limited to estimates based on sampling where the total number of ad impressions is estimated based on data from a subset of the potential ad recipients. The use of statistical based sampling, however, introduces sampling error. Where the sampling error is relatively small, the sampling based estimate can be useful for monitoring an ad's impression count. Where the sampling error is relatively large, however, the estimate may be of little value to advertisers.

[0007] Television, for example, is traditionally a broadcast medium. In order to measure the total number of viewers of a given program (or the advertisements therein), data from a subset of the potential viewer population is collected and extrapolated to create an estimate of the total number of viewers of the program. Historically, the number of networks was limited and the sampling error present in the viewer

estimates was relatively small. However, the large number of channels now available fragments the total audience such that the viewership of any one program can be a small fraction of the total viewer population. An estimate of the viewership in this case can have a sampling error that is so large that the estimate is not useful for determining viewership of the program and any ads included therein.

[0008] Lack of reliable data regarding a program's or network's viewership can lead to the network being forced to accept reduced rates for advertising time, or even to advertisers being reluctant to purchase advertising time on the network altogether. This can, in turn, lead to cancellation of a quality program or failure of a network, not because the programming was necessarily inferior and without an audience, but because traditional measurement techniques are incapable of providing accurate viewership or ad impression data that would monetize the value of advertising time for the niche programming.

[0009] The relatively small amount of profits generated by smaller networks, and the resulting lack of financial incentives to improve sampling techniques perpetuates the networks' and advertisers' dilemma. The smaller networks have advertising time to sell, but advertisers are reluctant to purchase the time due to the lack of verifiable return on the investment.

SUMMARY

[0010] This specification describes technologies relating to providing advertisement impression data.

[0011] In general, one aspect of the subject matter described in this specification can be embodied in methods that include the actions of estimating a number of advertisement impressions using a data sample and discounting the estimate to achieve a desired level of certainty that the resulting discounted estimate is less than the actual number of advertisement impressions. A sampling error model can be created for a data sampling and the discount required to achieve the desired certainty level can be determined using the model.

[0012] Particular embodiments of the subject matter described in this specification can be implemented to realize one or more of the following advantages. A metric estimator can be configured to supply discounted impression estimates to an advertisement system for use in advertiser billing. The discounted estimates can increase advertiser confidence that no overestimation exists, while permitting programmers to appropriately value and sell advertising time during difficult to measure programming.

[0013] The details of one or more embodiments of the subject matter described in this specification are set forth in the accompanying drawings and the description below. Other features, aspects, and advantages of the subject matter will become apparent from the description, the drawings, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] FIG. 1A is a flow chart of an example process for providing an advertisement impression count that, is adjusted to compensate for sampling error.

[0015] FIG. 1B is a flow chart of an example process for discounting an impression estimate that is based on sampled data.

[0016] FIG. 2 is a graph of an example sampling error distribution of estimated ad impressions based on a set of sampled data.

[0017] FIG. 3 is a graph showing the relationship between a desired certainty level and the discount required to achieve that level of certainty for an example set of data.

[0018] FIG. 4 is a graph showing the relationship between the discount level required to achieve a given certainty level, the sample size, and the number of consumers of a program for an example set of data.

[0019] FIG. 5 is a graph of per advertisement revenue versus sample size for an example set of data.

[0020] FIG. 6 is a graph of billable impressions versus sample size for an example set of data.

[0021] FIG. 7 is a block diagram of an example metric estimator of an advertisement system.

[0022] FIG. 8 is a graph of per advertisement revenue versus sample size which illustrates the incremental revenue between revenue amounts based on a first and second set of example sampled data.

[0023] FIG. 9 is a block diagram of an example environment for an advertising system.

[0024] Like reference numbers and designations in the various drawings indicate like elements.

DETAILED DESCRIPTION

[0025] In an implementation sampled data is used to generate an advertisement (“ad”) impression count that is adjusted to compensate for sampling error. In order to avoid overstated impression counts due to sampling error, impression estimates are discounted to a level such that it is known with a degree of certainty that the adjusted impression count is at or lower than the actual number of impressions received by the advertisement being measured.

[0026] FIG. 1A is a flow chart of an example process 100 for providing an advertisement impression count that is adjusted to compensate for sampling error. The sampled data is extrapolated to estimate a number of impressions received for a given advertisement (102). The estimated number of impressions is then discounted to ensure that the estimate is at or below an actual impression count (104). The initial estimate is discounted to an amount that is known with a desired degree of certainty to be at or lower than the actual number of impressions.

[0027] The process operates to avoid sampling error induced by overestimation of the number of ad Impressions. The discounted estimate can be used by advertisers as a lower threshold of impressions for a given advertisement, where the advertisers can be confident that the actual number of ad impressions is likely to be higher. The discounted estimate can also be used, for example, in billing advertising costs, such as in a pay-per-impression billing scheme. In this way advertisers can be confident that they are not being overcharged. The uncertainty present in relying on the sampled impression data is reduced and advertisers are more likely to purchase advertising time.

[0028] FIG. 1B is a flow chart of an example process 106 for discounting an impression estimate that is based on sampled data. A model of the sampling error present in the sampled data is created (108). In an implementation, the sampling error is modeled as a hypergeometric distribution. The amount that the estimate will be discounted is then determined based on the sampling error model. The estimate is discounted by an amount that, achieves a desired level of

certainty that the discounted estimate will be at or less than the actual number of impressions (110). The desired level of certainty can be a percentage of certainty that the resulting discounted estimate is at or less than an actual number of ad impressions. The discount is applied to the Initial estimate (112). In an implementation, the determined discount can be a percentage discount that is applied to the first estimate.

[0029] FIG. 2 is an example graph 200 of a sampling error distribution of estimated ad Impressions based on a set of sampled data. The distribution curve 202 peaks where the vertical line 204 crosses the distribution curve 202 at point 206. At this point the confidence of the estimate is greatest. This is the “best guess” estimate available of the number of ad impressions that the sampled data provides. However, due to sampling error, the model shows that it is possible that this estimate is higher than the actual number of impressions. There is a 50% chance that the actual number of impressions is higher than the estimate represented by point 206 and a 50% chance that the actual number of impressions is lower. There is an equal amount of area under the curve 202 to the right of vertical hue 204 as there is to the left of line 204.

[0030] Where a large range of sampling error exists, the actual number of impressions could vary by a large amount from the estimated number of impressions coinciding with point 206. In cases of potentially large sampling errors, a conservative estimate is desirable. Moving left on the graph for a distance ‘D’ 212 to the point on the distribution where the vertical line 208 crosses the distribution curve at point 210, yields a reduced value for the impression estimate. The chance that the actual number of impressions is higher than this reduced, or discounted, estimate (i.e. the estimate is low) has increased, and the chance that the actual number of impressions is lower (i.e. the estimate is high) has decreased. There is more area under the distribution curve 202 to the right of the vertical line 208 than there is to the left of it.

[0031] As the length of segment D is increased, the estimated impressions approach zero, the discount approaches 100%, and the certainty that the estimate is less than the actual number of impressions approaches 100%. As the length of D is decreased, the distribution estimate approaches its original value (the discount approaches zero), and the certainty that the estimate is less than, the actual number of impressions approaches 50%.

[0032] The amount of the discount depends on the level of certainty desired. The choice of certainly level is a compromise. On one hand, it is desirable to maximize advertiser confidence that impression data is not overstated. On the other hand, a programming provider would not want to sell itself short in its number of reported impressions.

[0033] The graphs of FIGS. 3-6 illustrate relationships between the desired certainty level, the required discount to achieve the desired certainty level, the sample size, and the popularity of a given program, such as a television or radio program.

[0034] FIG. 3 is a graph showing the relationship between, the desired certainty level and the discount required to achieve that level of certainty for an example set of data. The multiple graphed lines correspond to respective certainty levels. The highest discounts (vertical axis) are required to achieve a 99% certainty that a discounted number of impressions is less than the actual number of impressions, and the lowest discounts are required to achieve a 65% level of certainty for a given sample size. The graph also shows that as the size of the sample approaches the size of the total population,

which in this case is indicated as 100 million, the required discount approaches zero for each of the certainty level selections.

[0035] FIG. 4 is a graph showing the relationship between the discount level required to achieve a given certainty level, the sample size, and the number of consumers of a program for an example set of data. As in FIG. 3, the required discount approaches zero as the sample size approaches the total population (which for FIG. 4 is 13 million). The multiple graphed lines show that, for a given sample size, the less popular programs (those with lower ratings) require higher discounts.

[0036] FIG. 5 is a graph of per advertisement revenue versus sample size for an example data set. The example data set is for a program having a 0.01 percent rating out of a total viewer population of 13 million. The graph shows that an increase in the sample size (which in turn can increase the estimate satisfying the desired level of certainty) increases the revenue generated per ad.

[0037] FIG. 6 is a graph of billable impressions versus sample size for an example set of data. Three lines are shown for programs having respective levels of popularity. The graph shows that an increase in sample size for less popular programs has a more dramatic effect on the percentage of billable impressions than the same increase does for more popular programs.

[0038] Adjusting the estimated number of impressions by discounting the estimate in the manner described permits additional ad revenue to be generated from niche programming that would otherwise suffer from large sampling errors which could deter skeptical advertisers. The conservative impression data generated by the adjustment process increases advertiser confidence in impression data, which can encourage advertisers to purchase advertisements during niche programming. In cases where an advertiser is billed on a per impression basis, the adjustment can help prevent over billing. Furthermore, networks and operators (such as radio stations, cable television systems, and direct satellite broadcasters) are provided with an incentive to seek more thorough sampling data since larger samples can yield higher advertising revenues.

[0039] To facilitate the gathering of larger samples, networks and operators can enlist the help of third party data providers (a television ratings service, for example) who can, for example, share in the revenue increase made possible by the larger data samples. Where two or more sources of sampling data are available, the process permits an incremental revenue increase to be measured such that the increase in revenue attributable to larger data samples can be measured.

[0040] The disclosed process can be implemented in conjunction with an advertisement system configured to deliver relevant advertisements to consumers (e.g., subscribers, viewers, listeners, etc.) to facilitate provider (e.g., broadcast networks) monetization of current programming. For example, the disclosed process can be implemented in a metric estimator of an advertisement system.

[0041] FIG. 7 is a block diagram of a metric estimator 700 of an advertisement system. The metric estimator can be configured via hardware, software, or a combination thereof to implement the disclosed estimation adjustment process.

[0042] In a television advertising system, for example, the metric estimator 700 can be configured to generate an impression estimation based on sampled data. As described above, however, such sampling can result in errors, particularly during niche programming, e.g., programming with relatively

small viewership numbers. The metric estimator 700 can, for example, quantify the desired level of certainty for a metric, e.g., an impression count, and adjust the count such that the advertiser only pays for high-certainty impressions. Thus the advertiser may be more willing to purchase airtime during niche programming.

[0043] The sampling data 702 can include traditional television ratings information gathered by survey techniques and information gathered from device logs, such as television set top box logs.

[0044] In one implementation, the metric estimator 700 models sampling error 704 as a hypogeometric distribution and estimates the largest number of viewers such that there is a level of certainty, e.g., 90%, that the estimate is below the true value, given the expected sampling error. For example, advertisers may only be billed for the number of impressions that the metric estimator 700 determines with 90% certainty is less than, the number of actual impressions.

[0045] In one implementation, the hypogeometric distribution can be calculated for every sample. In another implementation, the hypogeometric distribution can be precomputed for a range of sample sizes. The precomputed data can be utilized to estimate the necessary discount required to achieve the 90% certainty level.

[0046] In one implementation, the metric estimator calculates a first impression estimate based on a first data sample (e.g., Nielsen television ratings data or a first set of set top box log information), and later calculates a second impression estimate based on a larger data sample (e.g., more set top box log information gathered since the first estimation). A budgeting system (not shown) calculates the incremental revenue generated by the second estimation. Any incremental revenue generated can be attributed to the second set of sample data. An upper limit for the sample size used by the metric estimator 400 can be determined as the sample size where the cost of obtaining a larger sample size exceeds the incremental revenue generated.

[0047] FIG. 8 is a graph of per advertisement revenue versus sample size which illustrates the incremental revenue between a first and second set of sampled data. The graphed data of the example relates to an impression data of an advertisement broadcast to a total population of 13 million subscribers where 1% of those subscribers watch or are at least tuned to the advertisement. An impression estimation calculation (adjusted to achieve a 90% certainty that the discounted estimate is less than the true number of impressions) based on a sample size of 400,000 viewers yields a per advertisement revenue of approximately \$61. A second estimate performed using a larger data sample (3 million viewers) yields a per advertisement revenue of approximately \$63.25. The incremental revenue attributable to the larger data sample is \$2.25. The first data sample can be, for example, set top box tuner log information received in a one day time window following the airing of the program or advertisement. The second data sample can be, for example, set top box tuner log information received during a four day time window following the airing of the program or advertisement, where the additional sample data is attributable to a greater number of set top boxes providing log data to the advertisement system through, for example, scheduled uploads.

[0048] In an implementation the metric estimator provides impression estimates to an advertising system. The advertising system can, for example, deliver relevant advertisements to viewers to facilitate operator monetization of program-

ming. Consumers, advertisers, providers and programmers are provided more relevant advertisements. A reporting process helps advertisers understand advertisement performance and to determine which advertisements are performing better than others to facilitate delivery of more relevant advertisements to consumers. Additionally, advertisers can, for example, more effectively test creative ideas and deliver advertising content that is more engaging to users and has greater effectiveness.

[0049] In another implementation, the advertising system can, for example, provide accurate and timely reporting for advertisers to improve the measurability of advertising. For example, in the television environment, set-top box based measurements with anonymized data can be used to report the number of advertisement impressions delivered and facilitate timely reporting so advertisers are informed on a timely basis as to when and where the advertisements will run or have run. The reporting capability further facilitates advertisers being better able to understand what advertisements are most effective with consumers. Aggregate statistics can also be reported from millions of set-top boxes, and the aggregate statistic can compliment other available data sources of information on television viewership, e.g., ratings systems.

[0050] In another implementation, the advertising system can, for example, facilitate expansion of the market by facilitating provider monetization of additional content. For example, relevant and effective advertisements for millions of television provider subscribers can be delivered locally and/or nationwide. Such relevant advertisements can more effectively monetize airtime, which can attract additional advertisers and create greater value for the network provider's advertisement inventory. Additionally, by measuring access of programmer inventory through subscriber set-top boxes, providers can more effectively monetize specialty channels that serve smaller audiences and for which audience data have been historically difficult to measure.

[0051] In another implementation, the advertising system can, for example, broaden the reach of an advertising medium, attracting new advertisers by making advertising more easily accessible through an automated advertising process. A video advertisement marketplace, for example, can facilitate identification of producers to assist prospective advertisers in the creation of video advertisements.

[0052] In another implementation, the advertising system can, for example, create efficiencies in the buying/selling process through an automated online marketplace. For example, in one implementation, the advertising process is automated from planning the campaign, uploading the advertisement and serving the advertisement. An auction model can be used to create pricing efficiencies for both buyers and sellers of advertising. Advertisers can benefit from efficiencies by paying only for delivered impressions and receiving the information the advertisers need to continually enhance the effectiveness of the advertiser's advertisements. The system can be implemented organically or can be implemented with third-party infrastructure partners.

[0053] In one implementation, the advertising system can measure set-top box activity using channel change records, which include tuned channels and associated time stamps. Multiple ways for advertisers to target audiences with traditional approaches, such as media, network/daypart, geographic and demographic, can also be offered.

[0054] However, targeting individual set-top boxes or viewers, for example, can be precluded. In one implementa-

tion, the set-top box and account information is anonymized before being received by the advertising system. Each set-top box can include an anonymous value and include channel change records. Account information can be anonymized and associated with location information that has no more granularity than a zip code. No demographic or psychographic information is associated with account information. No information gathered from other sources is associated with account information. The anonymization can be implemented to ensure the privacy of consumers, and new functionality can be introduced with a strong focus on maintaining consumer privacy.

[0055] The reporting information of a subscriber can be received by the subscriber's provider and the reporting information can be anonymized before being provided to the advertising system. As a result, the advertising system, in some implementations, does not have direct access to non-anonymized reporting information. The aggregation of reporting information and behavior thus does not target specific households.

[0056] The reporting information can, for example, be provided by third party providers to the advertising system. The reporting information can facilitate the delivery of effective and relevant advertisements while maintaining protections on subscriber privacy. The collection of such data can, for example, be managed to ensure compliance with the laws of local and national jurisdictions.

[0057] FIG. 9 is a block diagram of an example environment for an advertising system 900. In one implementation, the advertising system 900 can be implemented on one or more servers. However, in other implementations, any of the servers can be combined to perform multiple functions.

[0058] The advertising system 900 can include advertisement items (e.g., advertisements 902) and metadata 904. The advertisements 902 can, for example, include video advertisements, banner advertisements, overlay advertisements (e.g., logos), tickers (e.g., crawlers), voiceovers, etc. The metadata 904 can include audience data, impression data, performance data related to the advertisements 902, the advertisement metadata 904, filter data 906, etc. In an implementation, viewer information stored in the metadata 904 is anonymized.

[0059] In an implementation, the advertising system 900 can communicate with a software agent 910. The software agent 910 can, for example, be located at an advertising insertion location provider 920, e.g., a cable provider, a digital satellite television provider, newspaper, radio (e.g., satellite radio, terrestrial radio, etc.), cellular content provider, etc. The software agent 910 can, for example, optionally synchronize and cache advertisements, e.g., television videos, television banners, television overlays, streaming advertisements, tickers (e.g., crawlers), print, advertisements, voiceovers, etc., ahead of advertisement scheduling and that are provided by the advertising system 900.

[0060] In an implementation, the software agent 910 can read schedule requests, either in real time or ahead of time, and identify which schedule times the advertising system 900 has permission to fill with advertisements 902. In one implementation, the software agent 910 can request the advertising system 900 server to identify a relevant advertisement for an identified advertisement spot. The advertisement 902 can be deemed relevant based on the metadata 904 and other data. For example, an advertisement for extreme sporting equip-

ment may be selected for distribution during a sporting event for which the metadata identifies as a primary demographic 18-30 year old males.

[0061] The advertisement 902 can be scheduled based on the advertising system 900 response, and the advertisement 902 can be routed to advertisement equipment at the provider 920. The software agent 910 can read a status as to if and when the advertisement 902 actually released, and notify the advertising system 900 of the release. In another implementation, the software agent 110 can provide reporting data related to which of the consumers 922a-922n actually observed (e.g., saw, listened, read, selected, etc.) the advertisement 902. In an implementation, the reporting data can be anonymized.

[0062] In one implementation, the software agent 910 can monitor advertisement schedules that are not available for scheduling by the advertising system 900, and send the schedule data to the advertising system 900 for analysis of the advertisement market.

[0063] In an implementation, the advertising system 900 can accept requests from the software agent 910 for candidates for advertisement insertion and fetches potential candidate advertisements.

[0064] In one implementation, an advertisement filter 906 can perform one or more filtering processes to eliminate unwanted advertisements. In one implementation, the advertisement filter 106 can perform frequency capping that limits the scheduling of certain advertisements based on a certain amount of time since the advertisement was last distributed. In another implementation, an advertisement filter 906 applies competitive restrictions, where one advertisement cannot be placed near another advertisement of a competitor. Other filtering processes can also be used.

[0065] In an implementation, the advertising system 900 can identify advertisements based on account advertiser bids, budgets, and any quality metrics, e.g., conversions, viewer actions, sales numbers, etc., that have been collected.

[0066] The advertising system 900 can, for example, determine the prices for the advertisements, using one or more pricing processes. One exemplifying pricing process is a "second price pricer" where each advertiser pays the bid of the next highest advertisement. Other pricing processes can also be used, e.g., the estimate number of impressions multiplied by the price an advertiser has offered to pay for each impression.

[0067] The advertising system 900 can, for example, store in the metadata 904 relevant information related to advertisement scheduling requests and the results provided. This data may be used for internal analysis or the data can be exposed to advertisers.

[0068] The advertising system 900 can output the resulting advertisements as a response to a request from the software agent 910.

[0069] In an implementation, the advertising system 900 can facilitate the selection of advertisements for one particular "pod," e.g., a grouping of advertising space (e.g., an advertising break, a section of print, etc.). In another implementation, the advertising system 900 can facilitate selection of all available advertisements throughout a programming day across multiple networks (e.g., television, newspaper, radio, internet, etc.).

[0070] In an implementation, the advertisement data 902 and/or the metadata 904 can be stored on an advertisement candidate server. In various implementations, the advertise-

ment candidate server can be integrated into the advertisement system 900, or can be realized as a separate entity.

[0071] The advertisement candidate server can store all advertisements 902 that can be distributed, and any associated targeting data. The advertisement candidate server can handle requests from the television advertisement system 900, and can apply any targeting criteria before returning the advertisement(s) that can show for a particular request.

[0072] An example advertisement request and serving process can be as follows. The example process can be implemented in the system 900 of FIG. 9.

[0073] A request can be received for candidate television advertisements, for example, that can show In the USA/California/Bay Area/Mountain View, with a remoteRepositoryId of XX, for a partner provider of YY, on a television network ZZ, to be scheduled within the time window of Monday 2 PM-3 PM, most likely Monday 2:16 PM,

[0074] The candidate server can apply various targeting and filtering. For example, the advertisement must have been successfully synchronized to that remote repository; the advertisement can be targeting the location or a superset of the location where the advertisement will be showing; the advertiser/owner of an advertisement has adequate budget; and the advertisement is not paused, either by the advertiser or by the advertising system 900; the advertiser should not be considered fraudulent or delinquent; the partner provider, if agreed upon beforehand, has approved this advertisement for showing in these conditions; the advertisement is targeting this particular television network, content and/or time; the advertisement targets a television show which, through internal or third party data sources, corresponds to the given request; and the advertisement targets a demographic profile which, through internal or third party data sources, corresponds to the given request. Any combination of one or more such filter conditions can also be applied.

[0075] A verification process can accept distribution/status messages from the software agent 910 and update such information, e.g., update the metadata 104. This data may also be exposed to advertisers.

[0076] In an implementation, synchronizing of videos to the provider 920 can be facilitated, if needed. In one implementation, the software agent 910 periodically requests the advertising system 900 server for any new advertisements that need to be downloaded. For any such advertisements, the software agent 910 will initiate the download, and upon successful completion will notify the advertising system 900 of a successful download. The advertising system 900 can, for example, label the download with a particular ID that can be later used during scheduling to identify the scheduled advertisement.

[0077] Embodiments of the subject matter and the functional operations described in this specification can be implemented in digital electronic circuitry, or in computer software, firmware, or hardware. Including the structures disclosed in this specification and their structural equivalents, or in combinations of one or more of them. Embodiments of the subject matter described in this specification can be implemented as one or more computer program products, i.e., one or more modules of computer program instructions encoded on a tangible program carrier for execution by, or to control the operation of, data processing apparatus. The tangible program carrier can be a propagated signal or a computer readable medium. The propagated signal is an artificially generated signal, e.g., a machine generated electrical,

optical, or electromagnetic signal, that is generated to encode information for transmission to suitable receiver apparatus for execution by a computer. The computer readable medium can be a machine readable storage device, a machine readable storage substrate, a memory device, a composition of matter effecting a machine readable propagated signal, or a combination of one or more of them.

[0078] The term “data processing apparatus” encompasses all apparatus, devices, and machines for processing data, including by way of example a programmable processor, a computer, or multiple processors or computers. The apparatus can include, in addition to hardware, code that creates an execution environment for the computer program in question, e.g., code that constitutes processor firmware, a protocol stack, a database management system, an operating system, or a combination of one or more of them.

[0079] A computer program (also known as a program, software, software application, script, or code) can be written in any form of programming language, including compiled or interpreted languages, or declarative or procedural languages, and it can be deployed in any form, including as a stand alone program or as a module, component, subroutine, or other unit suitable for use in a computing environment. A computer program does not necessarily correspond to a file in a file system. A program can be stored in a portion of a file that holds other programs or data (e.g., one or more scripts stored in a markup language document), in a single file dedicated to the program in question, or in multiple coordinated files (e.g., files that store one or more modules, sub programs, or portions of code). A computer program can be deployed to be executed on one computer or on multiple computers that are located at one site or distributed across multiple sites and interconnected by a communication network.

[0080] The processes and logic flows described in this specification can be performed by one or more programmable processors executing one or more computer programs to perform functions by operating on input data and generating output. The processes and logic flows can also be performed by, and apparatus can also be implemented as, special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application specific integrated circuit).

[0081] Processors suitable for the execution of a computer program include, by way of example, both general and special purpose microprocessors, and any one or more processors of any kind of digital computer. Generally, a processor will receive instructions and data from a read only memory or a random access memory or both. The essential elements of a computer are a processor for performing instructions and one or more memory devices for storing instructions and data. Generally, a computer will also include, or be operatively coupled to receive data from or transfer data to, or both, one or more mass storage devices for storing data, e.g., magnetic, magneto optical disks, or optical disks. However, a computer need not have such devices. Moreover, a computer can be embedded in another device, e.g., a mobile telephone, a personal digital assistant (PDA), a mobile audio or video player, a game console, a Global Positioning System (GPS) receiver, to name just a few.

[0082] Computer readable media suitable for storing computer program instructions and data include all forms of non volatile memory, media and memory devices, including by way of example semiconductor memory devices, e.g., EPROM, EEPROM, and flash memory devices; magnetic disks, e.g., internal hard disks or removable disks; magneto

optical disks; and CD ROM and DVD ROM disks. The processor and the memory can be supplemented by, or incorporated in, special purpose logic circuitry.

[0083] To provide for interaction with a user, embodiments of the subject matter described in this specification can be implemented on a computer having a display device, e.g., a CRT (cathode ray tube) or LCD (liquid crystal display) monitor, for displaying information to the user and a keyboard and a pointing device, e.g., a mouse or a trackball, by which the user can provide input to the computer. Other kinds of devices can be used to provide for interaction with a user as well; for example, feedback provided to the user can be any form of sensory feedback, e.g., visual feedback, auditory feedback, or tactile feedback; and input from the user can be received in any form, including acoustic, speech, or tactile input.

[0084] While this specification contains many specific implementation details, these should not be construed as limitations on the scope of any invention or of what may be claimed, but rather as descriptions of features that maybe specific to particular embodiments of particular inventions. Certain features that are described in this specification in the context of separate embodiments can also be implemented in combination in a single embodiment. Conversely, various features that are described in the context of a single embodiment can also be implemented in multiple embodiments separately or in any suitable subcombination. Moreover, although features may be described above as acting in certain combinations and even initially claimed as such, one or more features from a claimed combination can in some cases be excised from the combination, and the claimed combination may be directed to a subcombination or variation of a subcombination.

[0085] Similarly, while operations are depicted in the drawings in a particular order, this should not be understood as requiring that such operations be performed in the particular order shown or in sequential order, or that all illustrated operations be performed, to achieve desirable results. In certain circumstances, multitasking and parallel processing may be advantageous. Moreover, the separation of various system components in the embodiments described above should not be understood as requiring such separation in all embodiments, and it should be understood that the described program components and systems can generally be integrated together in a single software product or packaged into multiple software products.

[0086] Particular embodiments of the subject matter described in this specification have been described. Other embodiments are within the scope of the following claims. For example, the actions recited in the claims can be performed in a different order and still achieve desirable results. As one example, the processes depicted in the accompanying figures do not necessarily require the particular order shown, or sequential order, to achieve desirable results. In certain implementations, multitasking and parallel processing may be advantageous.

1. A method, comprising:

estimating a number of advertisement impressions using a data sample; and
discounting the estimate to achieve a desired certainty level that a resulting first discounted estimate is at or less than an actual number of advertisement impressions.

2. The method of claim 1 wherein the desired level of certainty is in a range of 90 to 95 percent.

- 3. The method of claim 1 wherein discounting the estimate comprises:
 - creating a sampling error model for the data sample; and
 - determining the discount required to achieve the desired level of certainty based on the sampling error model.
- 4. The method of claim 3 wherein discounting the estimate further comprises;
 - applying the determined discount to the estimated number of advertisement impressions.
- 5. The method of claim 1 further comprising:
 - reporting the discounted estimate to an advertiser responsible for the advertisement.
- 6. The method of claim 5 further comprising:
 - billing the advertiser based on the discounted estimate.
- 7. The method of claim 1 wherein the data sample comprises survey results.
- 8. The method of claim 1 wherein the data sample comprises device logs.
- 9. The method of claim 1 further comprising:
 - creating a second discounted estimate using a second data sample that is larger than the first;
 - billing an advertiser an amount based on the second discounted estimate;
 - calculating a difference between the amount billed and an amount based on the first discounted estimate.
- 10. The method of claim 9 further comprising:
 - providing a portion of the difference to a supplier of the second data sample.
- 11. The method of claim 3 wherein the sampling error model is a hypergeometric distribution.
- 12. A system comprising:
 - a metric estimator configured to:
 - receive sampling data;
 - receive sampling error information for the sampling data;
 - estimate a number of advertisement impressions based on the sampling data to create a first estimate;
 - model the sampling error of the sampling data using the sampling error information; and
 - discount the first estimate by an amount determined using the sampling error model to achieve a desired certainty level that the resulting discounted estimate is at or less than an actual number of advertisement impressions.
- 13. The system of claim 12 wherein the sampling error model is a hypergeometric distribution.

- 14. The system of claim 12 wherein the desired certainty level that the discounted estimate is at or less than the actual number of advertisement impressions is in a range of 90 to 95 percent.
- 15. The system of claim 12 wherein the sampling data comprises survey results.
- 16. The system of claim 12 wherein the sampling data comprises device log data.
- 17. The system of claim 12 wherein the metric estimator is further configured to provide the discounted estimate to an advertisement system for billing an advertiser.
- 18. The system of claim 12 wherein the metric estimator is further configured to:
 - receive second sampling data;
 - receive second sampling error information for the second sampling data;
 - estimate a number of advertisement impressions based on the second sampling data to create a second estimate;
 - model the second sampling error of the second sampling data using the second sampling error information;
 - discount the second estimate by an amount determined using the second sampling error model to achieve the desired certainty level that the resulting second discounted estimate is at or less than the actual number of advertisement impressions; and
 - calculate the difference between the first discounted estimate and the second discounted estimate.
- 19. A system configured to:
 - accept an advertisement from an advertiser;
 - insert the advertisement into a commercial break of a broadcast program;
 - calculate an initial estimate using sampled data, the number of impressions received by the advertisement during the broadcast;
 - calculate a discounted estimate using a sampling error model, the discounted estimate having a calculated degree of certainty that it is at or lower than the actual number of impressions received by the advertisement during the broadcast; and
 - bill the advertiser a per impression rate based on the discounted estimate.
- 20. The system of claim 19 wherein the sampling error model is a hypergeometric distribution.

* * * * *