



US 20220398476A1

(19) **United States**(12) **Patent Application Publication**
Heikell(10) **Pub. No.: US 2022/0398476 A1**(43) **Pub. Date: Dec. 15, 2022**(54) **SYSTEMS, METHODS AND APPARATUS
FOR EVALUATING STATUS OF
COMPUTING DEVICE USER**(52) **U.S. Cl.**CPC **G06N 7/005** (2013.01); **G06F 21/50**
(2013.01); **G06F 21/316** (2013.01); **G06F**
2221/2133 (2013.01)(71) Applicant: **Nobots LLC**, Renton, WA (US)(72) Inventor: **Timothy P. Heikell**, Renton, WA (US)(21) Appl. No.: **17/882,082**(22) Filed: **Aug. 5, 2022****Related U.S. Application Data**

(60) Continuation of application No. 16/578,823, filed on Sep. 23, 2019, which is a division of application No. 15/457,099, filed on Mar. 13, 2017, now Pat. No. 10,423,885, which is a continuation of application No. 12/313,502, filed on Nov. 19, 2008, now Pat. No. 9,595,008.

(60) Provisional application No. 61/003,743, filed on Nov. 19, 2007.

Publication Classification(51) **Int. Cl.****G06N 7/00** (2006.01)**G06F 21/50** (2006.01)**G06F 21/31** (2006.01)(57) **ABSTRACT**

Methods, systems and apparatus, e.g., non-transitory machine readable medium having stored thereon instructions, for providing a value indicating a judgment of whether a human or an autonomic computer application operating a client computer is interacting with a server, by operations comprising: receive data from the client computer including active data relating to interactions of the client computer with a website of the server and passive data of the client computer; analyze received data from the client computer for indication of a human operating the client computer, wherein the analyzing includes analyzing the received data in conjunction with model data based on human interactions from a prior session with the same website and developing a first analysis value; and provide a request for further data from the client computer when the first analysis value fails to meet the first analysis predetermined criteria.

Nobots LLC - Ex. 2007
IPR2026-00429
Google LLC v. Nobots

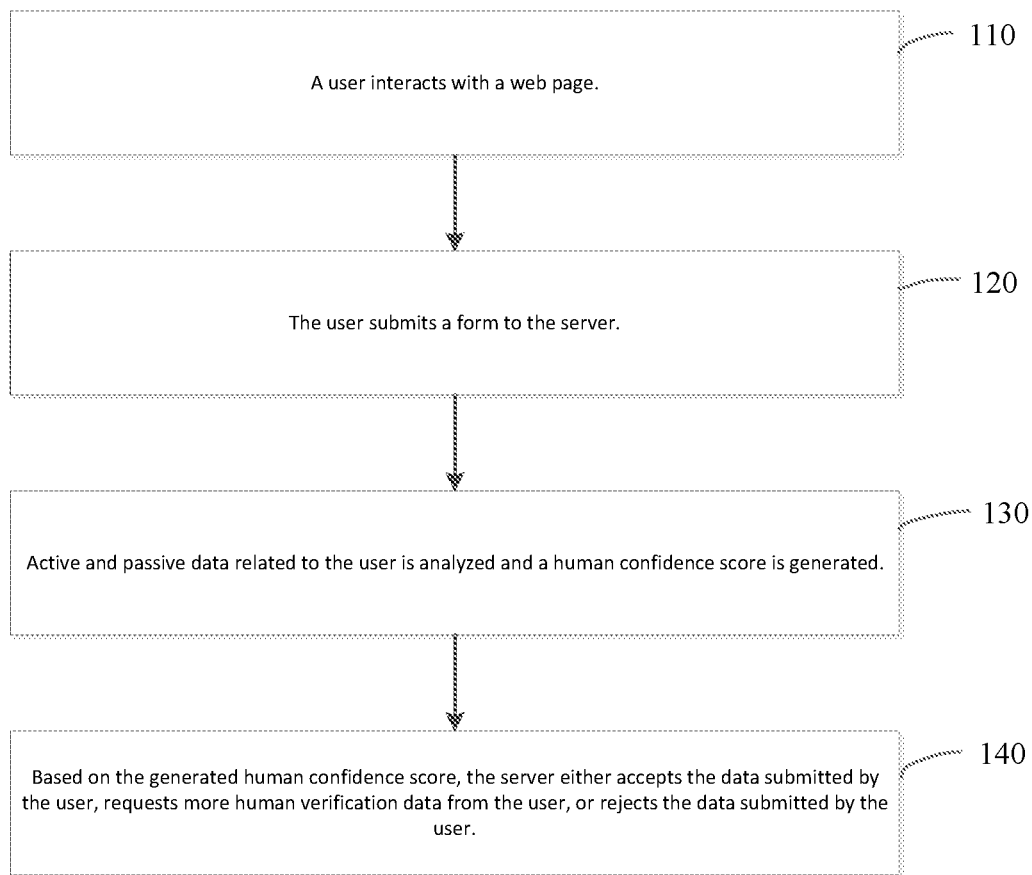


Fig 1

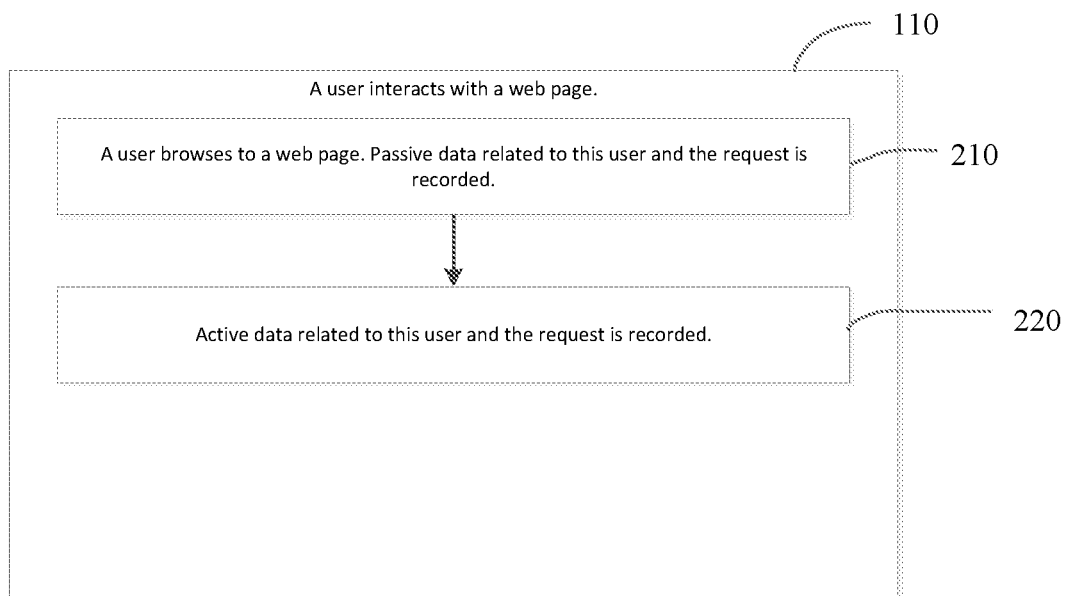


Fig 2

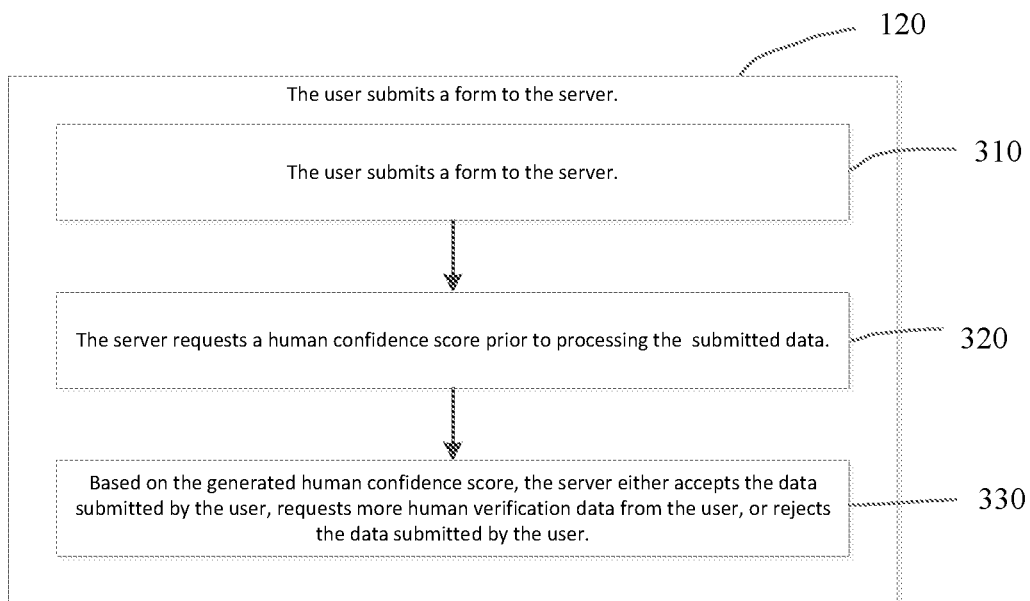


Fig 3

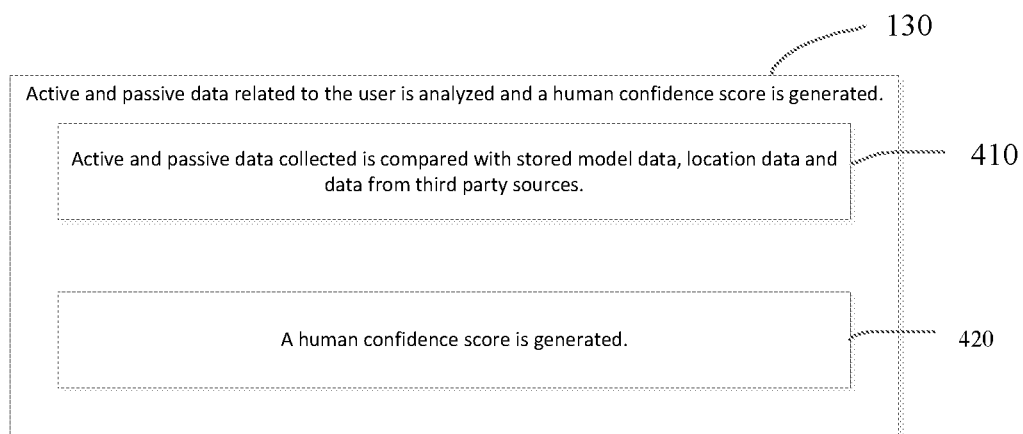


Fig 4

SYSTEMS, METHODS AND APPARATUS FOR EVALUATING STATUS OF COMPUTING DEVICE USER

[0001] This application is a continuation application of and claims priority to application Ser. No. 16/578,823 filed Sep. 23, 2019, which is a divisional application of application Ser. No. 15/457,099 filed Mar. 13, 2017 and issued as U.S. Pat. No. 10,423,885 on Sep. 24, 2019, which claims priority to application Ser. No. 12/313,502 filed Nov. 19, 2008 and issued as U.S. Pat. No. 9,595,008 on Mar. 14, 2017, which claims priority to provisional application Ser. No. 61/003,743 filed Nov. 19, 2007, all of which are incorporated herein by reference.

BACKGROUND

[0002] The Internet is a fantastic tool for constructive web sites to gather users for a common purpose; however, the Internet is also a fantastic tool for abuse of these same web sites. People who want to take advantage of websites do so by creating automated programs employing various algorithms and routines (hereinafter “bots”) that create fictitious accounts or access content for a multitude of reasons.

[0003] In an effort to block these bots, builders of web sites have created a variety of tests to determine if the user is a bot or if the user is a human. Initial efforts required a user to simply enter an alphanumeric string into an input field. However, as character recognition engines became more available, such “tests” became easily defeated. What was needed was a more robust form of test—one that couldn’t be easily defeated.

[0004] Carnegie Mellon University coined the term “CAPTCHA” (Completely Automated Public Turing test to tell Computers and Humans Apart) for these types of tests. A common type of CAPTCHA requires that the user type the letters, digits or characters of a distorted image appearing on the screen. The objective is to create an image that a bot cannot easily parse but that is discernable by a human. Such efforts have been successful in preventing non-adaptive software from recognizing the imaged characters, but people intent on abusing these sites have designed ways to circumvent the CAPTCHA, such as through specially tuned character recognition programs. A brief survey of the Internet will reveal many resources that describe how to tune and/or use character recognition to decipher CAPTCHA including aiCaptcha, Simon Fraser University and PWNtcha.

[0005] The result of the foregoing is that while CAPTCHAs are becoming increasingly more difficult for bots, they are also becoming more difficult and/or burdensome for human users. In certain instances, the desire to defeat the bots has resulted in images that are so distorted that some human users cannot decipher the images. This is particularly true with users having a visual deficiency or impairment. As a partial solution to this escalation of perception difficulty, some web sites have begun adding a link to a sound file that will speak the characters, but these sound files are also being drastically distorted to protect against being discerned by bots through speech pattern matching algorithms. Other web sites like Facebook.com, have gone so far as to adopt a practice requiring deciphering two distorted word images to increase the complexity for bots. While perhaps achieving the stated objective, the collateral effect is to exacerbate the existing burden to human users.

[0006] Current CAPTCHA technology is visual or auditory in nature, requiring the human user to answer a test that should be simple to most humans but difficult for non-humans, e.g., bots. Visual CAPTCHA using distorted images is widely used as the primary test by nearly every top Internet site including Yahoo, Google, You Tube, Microsoft’s Live ID, MySpace, Facebook, Wikipedia, Craigs List. By using solely visual testing criteria, nearly all users will be able to invoke the requested action; not all users have functioning audio equipment or environments such as libraries may not permit such use.

[0007] A positive user experience is critical to the success and increased popularity of a given website. Designers of web sites go to great lengths to ensure their website is as user friendly as possible. Carnegie Mellon University estimates that 60 million CAPTCHA tests are deciphered every day and with an average time spent of 10 seconds, requiring a total of 150,000 hours of work spent every day trying to protect web sites from bots. Reducing or eliminating the requirement of a user having to decipher CAPTCHA is one more way websites can create a more positive user experience for their visitors and minimize opportunity costs.

SUMMARY OF THE INVENTION

[0008] The invention is generally directed to methods, systems and apparatus for assessing the likely user status of a computing device interacting with a server where computing device is in bi-directional operative communication with the server wherein the status is one of a human operator or a computer executable program (also referred to herein as a “bot”). This assessment comprises comparing acquired and/or available data relating to the operation of the computing device to suitable models embodying human user derived data (model data). In most embodiments, the comparison yields a probability value as to one of the status states **140**, **330**, which then may be used by a program or administrator of the server to permit or deny access and/or operation to the computing device. Because many of the invention embodiments provide a probability result as opposed to a binary result, the invention embodiments avoid the “there is only one right answer” phenomena inherent in prior art CAPTCHA tests. In other words, rather than placing the burden of proof on the user for functionality/access, which if the user is a human invokes the negative consequences of conventional CAPTCHA tests as previously described, the burden is shifted to the server side of the equation.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] The detailed description is described with reference to the accompanying figures. The use of the same reference numbers in different figures indicates similar or identical components or features.

[0010] FIG. 1 illustrates an overview of the process described in this disclosure.

[0011] FIG. 2 illustrates in more detail the first step **110** of FIG. 1 (a user interacts with a web page).

[0012] FIG. 3 illustrates in more detail the second step **120** of FIG. 1 (the user submits a form to the server).

[0013] FIG. 4 illustrates in more detail the third step **130** of FIG. 1 (active and passive data related to the user is analyzed and a human confidence score is generated).

DETAILED DESCRIPTION

[0014] As used herein, “model data”, its equivalents and verb forms comprises data indicative of human interaction with a computing environment and that can be received by a computing device that is physically remote from the sample computing environment and equivalents. Model data comprises two main categories: active model data **220** and passive model data **210**. Active model data comprises data acquired from a computing device user’s interactions therewith and within the computing environment where such data is not normally stored (logged) or transmitted to a remote location. Such model data includes, without limitation, pointing device vector movements and/or cadence, key stroke combinations and/or cadence, time differentials between stimulus (e.g., display of dialog box, radio button, form field, etc., and/or generation of sound) and user response (e.g., input into dialog box, selection of radio button, completion of form field, new page display request rates, etc., and/or input response to sound), and similar metrics. Generally, such data must be monitored and stored **210, 220** by a program operative on the computing device, which makes the data available to another program, preferably on a server **320**, or actively transmits such data to a server. Passive model data comprises data available from a computing device user’s interactions therewith and within the computing environment where such data is normally stored (logged) or transmitted to a remote location. Such model data includes, without limitation, browser cookies, destination IP histories, originating IP address, originating IP address traffic data, originating IP address physical location, third party data regarding abusers (including originating IP addresses and physical locations), etc.

[0015] Also as used herein, the term “available data”, its equivalents and verb forms comprises data associated with a computing device’s operation and its interaction with a computing environment, such as the Internet, that is generally recorded within the computing device and/or by other devices that have been affected by the computing device’s operation this is also a type of passive data; the term “acquired data”, its equivalents and verb forms comprises data associated with a computing device’s operation and its interaction with a computing environment, such as the Internet, that is generally not recorded within the computing device and/or by other devices that have been affected by the computing device’s operation, but at least some data of which has/have been recorded and/or transmitted to a remote location, such as a server—this is a type of active data.

[0016] In addition to the foregoing, the term “issued data”, its equivalents and verb forms comprises data generated by a server or other computing device that is not the same as the computing device for which the assessment as to user status is being performed; “monitored data”, its equivalents and verb forms comprises active or passive data, whether available or acquired, obtained from the computing device, or as a result of its external interactions, after the generation of issued data; “interest data”, its equivalents and verb forms comprises active or passive data, whether available or acquired, that correlates to any data within model data, whether obtained prior to or after the generation of issued data. Thus, interest data includes time independent available data and acquired data, unless qualified differently.

[0017] With the foregoing definitions in mind, operation of the various invention embodiments can be better understood. In a first series of embodiments, a comparison

between interest data, acquired prior to delivery of issued data to the client computing device, and model data is performed to ascertain the likely status of the client computing device, i.e., human user or bot **130, 420**. In a second series of embodiments, a comparison between monitored data, by definition acquired after delivery of issued data to the client computing device, and model data is performed to ascertain the likely status of the client computing device, i.e., human user or bot **130, 420**. In both series of embodiments, acquired and/or available data may be used for comparison with suitable model data. The recited comparisons can take place locally on the computing device, remotely on the originating server, or on a server dedicated to performing such actions and for which subscriptions may be offered in conjunction with methods for providing services according to the methods, apparatus and systems embodiments described herein.

[0018] While available data represents data that is readily harvestable by query, for example, from the computing device or the computing environment in which the device operates, acquired data requires some form of information capture means. In the various embodiments described herein, the computing device is caused to monitor and retain certain data useful as acquired data for comparison purposes. Such monitoring and retaining means for acquiring data from the computing device comprises, without limitation, modification of (an) existing program(s) (e.g., such means are included in available browsers), a covert program (e.g., many malware applications log keystrokes and periodically pass them to remote servers for malicious purposes; similar technology can be used to exploit necessary aspects of the invention embodiments), or a servlet/Java applet. If user privacy is a concern, the monitoring and retaining means can remain dormant until activated by, for example, an enabled web site **110**.

[0019] The monitoring and retaining means may also enable transmission of some or all retained data **410**, in encrypted or unencrypted form, as may be desired for privacy and security purposes, and/or merely retain the data until requested from, for example, the server, at which time some or all data may be transmitted **120, 310**. As described above with reference to the comparison actions **130, 410**, such receiving and/or polling actions can be carried out remotely on the originating server or on a server dedicated to performing such actions, if not performed locally on the computing device.

[0020] From the foregoing, it can be seen that implementation of the invention embodiments can be accomplished exclusively from the server side; it is not necessary to distribute or install in the conventional sense client side software. Existing available browsers and operating systems provide the means necessary to temporarily install logging code, if such is elected. Moreover, the methods, and associated systems and apparatus, described herein are highly transparent to the user, thereby achieving an objective of enhancing the user’s experience of a web site employing bot assessment protocols.

DESCRIPTION OF AN INVENTION EMBODIMENT

[0021] A primary objective of bot creation is to autonomously access data and/or functionality of a target server as quickly as possible. By assessing user biometrics having a time domain, the time variable becomes a necessary com-

ponent to accessing the data and/or functionality of the server. Because such assessment has heretofore been absent as a valid CAPTCHA marker of a human user, and more importantly because proper data input would necessarily slow the process, the likelihood of bot penetration has been significantly reduced.

[0022] An embodiment of the invention employs a first layer of testing that simply checks if there were valid mouse movements and/or key strokes inputted by the user of a computing device that is attempting to access a server resource “protected” from bots. This basic “if-then” check is essentially without overhead since there are no computations being carried out.

[0023] Checking for the existence of the target activity therefore represents a first pass evaluation; if the bot is not programmed to include pseudo biometric data, further access is denied. In other words, if no activity is recorded there is a very high probability that the user is actually a bot.

[0024] A fundamental premise of robust biometrics is that a given dataset for each person is unique. Therefore, if the dataset is sufficiently robust, it is impossible to have duplicative input data unless the input data was derived from a machine. Exploiting this premise allows a second level knockout assessment to deny user access if the input data exactly (or statistically sufficiently) matches previously recorded data. Of course, the skilled practitioner employing this method can select (either explicitly or via programming) sample points of a dataset for comparison as opposed to all data, thereby reducing computational overhead and storage issues. Alternatively, if samples are used, an exact match could then invoke a more intensive comparison with the same stored datasets, where again access can be denied when an exact or statistically sufficient match is found.

[0025] In the foregoing two assessments, an object has been to ferret out bots in an efficient and low overhead manner by exploiting intrinsic design limitations. However, it is possible that a bot designer could spoof these assessment means by, for example, running many bots in parallel wherein intrinsic delays in CPU processing and bandwidth would introduce inherent time delays associated with the very inputs being assessed. Therefore, more robust assessment means may be employed to ascertain the presence of a bot.

[0026] In robust embodiments of the invention, a third layer of testing may be employed that compares recorded pointer movements and key strokes to previously recorded activity for a given input page that was knowingly created by humans. Thus, as input data is collected for a given page, patterns will emerge that are unique to human activity. Subsequently recorded activity that is inconsistent with these patterns would indicate the potential that the user is a bot. Access could then be denied, or further CAPTCHA tests presented. Alternatively, access could be granted since no lock is pick proof and an object of the invention embodiments is to minimize user exposure to CAPTCHA tests.

1. A non-transitory machine readable medium having stored thereon instructions for performing a method for generating a value indicative of a likelihood that a human or autonomic computer application is operating a client computer attempting access to a protected page of a website, which when executed by at least one machine, causes the at least one machine to perform operations comprising:

provide data collection data that causes the client computer to collect active data relating to interactions of the client computer with a given page of the website;
receive at least some of the collected active data from the client computer and at least some passive data from the client computer;

perform a first analysis comprising analyzing at least some of the received active data and at least some of the received passive data in conjunction with model data based on human interactions from a prior session with the same website;

generate a first analysis value based on the first analysis, wherein the client computer is allowed to access the protected page without requiring an operator of the client computer to decipher a CAPTCHA test if the first analysis value meets a first analysis predetermined criteria;

provide a CAPTCHA test to the client computer when the first analysis value fails to meet the first analysis predetermined criteria;

receive a response of the client computer to the CAPTCHA test;

perform a second analysis comprising assessing accuracy of the response of the client computer to the CAPTCHA test; and

generate a second analysis value based on the second analysis, wherein the client computer is granted access to the protected page based on the second analysis value meeting a second analysis predetermined criteria.

2. The non-transitory machine readable medium of claim 1, further comprising instructions, when executed by the at least one machine, that cause the at least one machine to:

provide an additional CAPTCHA test to the client computer when the second analysis value fails to meet the second analysis predetermined criteria;

receive a response of the client computer to the additional CAPTCHA test;

perform a third analysis comprising assessing the accuracy of the client computer's response to the additional CAPTCHA test; and

generate a third analysis value based on the third analysis, wherein the client computer is granted access to the protected page based on the third analysis value meeting a third analysis predetermined criteria.

3. The non-transitory machine readable medium of claim 1, wherein the data collection data comprises JavaScript code that causes a browser of the client computer to collect the active data.

4. The non-transitory machine readable medium of claim 3, wherein the received passive data comprises at least one browser cookie and at least one Internet Protocol (IP) address associated with the client computer, and wherein the step of analyzing at least some of the received active data and at least some of the received passive data in conjunction with model data further includes analyzing the at least one browser cookie and the at least one IP address in conjunction with the model data.

5. The non-transitory machine readable medium of claim 4, wherein analyzing at least some of the received active data and at least some of the received passive data in conjunction with model data includes analyzing manual device input activity of the client computer in conjunction with the model data based on human interaction with the same given page.

6. The non-transitory machine readable medium of claim 5, wherein the manual device input activity comprises mouse movement over a CAPTCHA test prompt.

7. The non-transitory machine readable medium of claim 5, wherein the human interaction from a prior session with the same website includes human interaction of another client computer from the prior session with the same given page.

8. The non-transitory machine readable medium of claim 7, wherein the received active data comprises keyboard activity of the client computer on the given page and wherein the step of analyzing at least some of the received active data and at least some of the received passive data in conjunction with model data further includes analyzing the keyboard activity in conjunction with the model data.

9. The non-transitory machine readable medium of claim 1, wherein when the first analysis value meets a first analysis predetermined criteria, the client computer is not presented with a CAPTCHA test prior to the client computer being granted access to the protected page.

10. The non-transitory machine readable medium of claim 1, further comprising instructions, when executed by the at least one machine, that cause the at least one machine to decrypt the received active data and the received passive data.

11. The non-transitory machine readable medium of claim 1, wherein the model data is based on human interaction from a prior session with the same given page.

12. The non-transitory machine readable medium of claim 1, wherein the ability of the client computer to collect active data is only active while the client computer is visiting a web page enabled with data collection data.

13. The non-transitory machine readable medium of claim 1, further comprising instructions, when executed by the at least one machine, cause the at least one machine to receive the first or second analysis predetermined criteria from an operator or administrator of the website.

14. The non-transitory machine readable medium of claim 1, further comprising instructions, when executed by the at least one machine, that cause the at least one machine to receive from an operator or administrator of the website a number of CAPTCHA tests that will be provided to a client computer after the second analysis value fails to meet the second analysis predetermined criteria.

15. The non-transitory machine readable medium of claim 1, wherein the step of analyzing at least some of the received active data and at least some of the received passive data in conjunction with model data further comprises analyzing mouse and keyboard activity of the client computer in conjunction with mouse and keyboard activity from human interactions from a prior session with the same website.

16. A non-transitory machine readable medium having stored thereon instructions for performing a method for providing a value indicating a judgment of whether a human or an autonomic computer application operating a client computer is interacting with a server, which when executed by at least one machine, causes the at least one machine to perform operations comprising:

provide data collection data requesting active and passive data of the client computer;

receive active and passive data responsive to the data collection data, the received active data relating to interactions of the client computer with a website of the server;

analyze the received active data and the received passive data for indication of a human or an autonomic computer application operating the client computer, wherein the analyzing includes analyzing the received active data and the received passive data in conjunction with model data based on human interactions from a prior session with the same website and developing a first analysis value; and

provide a request for further data from the client computer when the first analysis value fails to meet the first analysis predetermined criteria.

17. The non-transitory machine readable medium of claim 16, further comprising instructions, when executed by the at least one machine, that cause the at least one machine to:

receive further data from the client computer in response to the request for further data;

perform a at second level of analyzing on the received further data; and

provide a second analysis value based on the second level of analyzing.

18. The non-transitory machine readable medium of claim 17, wherein the received further data includes the response of the client computer to a CAPTCHA test provided to the client computer.

19. The non-transitory machine readable medium of claim 16, wherein the first analysis value failing to meet the first analysis predetermined criteria indicates that the client computer should not be allowed access to a protected page of the website.

20. The non-transitory machine readable medium of claim 16, wherein the received active data includes manual input device activity and the received passive data includes at least one browser cookie and at least one IP address, and the analyzing includes analyzing the at least one browser cookie, the at least one IP address, and the manual input device activity in conjunction with model data based on human interactions from a prior session with the same website.

21. The non-transitory machine readable medium of claim 16, further comprising instructions, when executed by the at least one machine, that cause the at least one machine to decrypt the received active data and the received passive data.

22. A non-transitory machine readable medium having stored thereon instructions for performing a method for providing a value indicating a judgment of whether a human or an autonomic computer application operating a client computer is interacting with a server, which when executed by at least one machine, causes the at least one machine to perform operations comprising:

provide instructions that allow client computer interactions with a website of the server to be remotely monitored;

receive monitored data originating with the client computer, the monitored data comprising active data relating to interactions of the client computer with the website and passive data of the client computer;

analyze the received monitored data for indication of a human or an autonomic computer application operating the client computer, the analyzing including:

perform a first level of analyzing on the received monitored data for any indication that a human may be operating the client computer to obtain a first-level analysis;

perform a second level of analyzing to obtain a second-level analysis when the first-level analysis does not meet a first-level analysis predetermined criteria, the second level of analyzing comprising analyzing the received monitored data in conjunction with previously recorded data relating to human interactions from a prior session with the same website; and

provide a second-level analysis value when the second-level analysis meets a second-level analysis predetermined criteria.

23. The non-transitory machine readable medium of claim **22**, the analyzing further comprising instructions, when executed by the at least one machine, that cause the at least one machine to:

provide a request for further data from the client computer when the second-level analysis does not meet the second level analysis predetermined criteria.

24. The non-transitory machine readable medium of claim **23**, the analyzing further comprising instructions, when executed by the at least one machine, that cause the at least one machine to:

receive further data from the client computer in response to the request for further data;

perform a third-level of analyzing on the received further data; and

provide a third-level analysis value after completion of the third-level of analyzing.

25. The non-transitory machine readable medium of claim **24**, wherein the request for further data includes a CAPTCHA test.

26. The non-transitory machine readable medium of claim **22**, wherein the received monitored data includes at least one browser cookie and at least one IP address.

27. The non-transitory machine readable medium of claim **26**, wherein the received monitored data includes manual device input activity.

28. The non-transitory machine readable medium of claim **27**, wherein the manual device input activity comprises mouse movement and key strokes.

29. The non-transitory machine readable medium of claim **22**, further comprising instructions, when executed by the at least one machine, that cause the at least one machine to decrypt the received monitored data.

30. The non-transitory machine readable medium of claim **22**, wherein the monitored data comprises active data relating to interactions of the client computer with a given page of the website and passive data of the client computer, and the recorded data relates to human interaction from a prior session with the same given page.

* * * * *