actually be dealing with a more complex conversation requiring more time and in-depth analysis. At best, some systems may add or subtract a predetermined slot fraction to conversations flagged as especially simple or especially difficult. Such predetermined slot fractions cannot, however, identify the predicted workload with any more precision and remain unable to distinguish between different levels of difficulty.

[0005] Evaluating conversations so as to ascertain a predicted workload (e.g., in terms of time and engagement level) may be complicated, however, as no conversation is conducted alike. Different individuals of different backgrounds use language differently, and there is, therefore no standardized way to evaluate their respective conversations.

SUMMARY

[0005a] According to a first aspect of the present disclosure, there is provided a method including: storing a workload model for a set of analysis elements associated with user communications, wherein the set of analysis elements are associated with a measurement of workload; receiving an incoming request for a user communication, the incoming request including information regarding one or more analysis elements of the set of analysis elements; identifying a workload measurement for the user communication based on comparing the incoming request to the workload model; identifying an agent to handle the user communication, wherein the agent is identified based on a current workload capacity for the agent and the workload model indicating the agent as having the current workload capacity to handle a workload indicated by the workload measurement; activating a communication slot for the agent, the communication slot defined based on the workload measurement; routing the incoming request to the agent; and setting the current workload capacity for the agent based on the workload measurement.

[0005b] According to a second aspect of the present disclosure, there is provided a system including: a memory that stores a workload model for a set of analysis elements associated with user communications, wherein the set of analysis elements are associated with a measurement of workload; and one or more processors coupled to the memory and configured to perform operations including: receiving an incoming request for a user communication, the incoming request including information regarding one or more analysis elements of the set of analysis elements; identifying a workload measurement for the user communication based on comparing the incoming request to the workload model; identifying an agent to handle the user communication, wherein the agent is identified based on a current workload capacity for the agent and the workload model indicating the agent as having the current workload capacity to handle a workload indicated by the workload measurement; activating a communication slot for the agent, the communication slot defined based on the workload measurement; initiating routing of the incoming request to the agent; and setting the current workload capacity for the agent based on the workload measurement.

[0005c] According to a third aspect of the present disclosure, there is provided a non-transitory computer readable medium including instructions that, when executed by one or more processors of a device, cause the device to perform operations for workload capacity routing including: storing a workload model for a set of analysis elements associated with user communications, wherein the set of analysis elements are associated with a measurement of workload; receiving an incoming request for a user communication, the incoming request including information regarding one or more analysis elements of the set of analysis elements; identifying a workload measurement for the user communication based on comparing the incoming request to the workload model; identifying an agent to handle the user communication, wherein the agent is identified based on a current workload capacity for the agent and the workload model indicating the agent as having the current workload capacity to handle a workload indicated by the workload measurement; activating a communication slot for the agent, the communication slot defined based on the workload measurement; routing the incoming request to the agent; and setting the current workload capacity for the agent based on the workload measurement.

[0006] The term example and like terms are intended to refer broadly to all of the subject matter of this disclosure and the claims below. Statements containing these terms should be understood not to limit the subject matter described herein or to limit the meaning or scope of the claims below. Examples of the present disclosure covered herein are defined by the claims below, not this summary. This summary is a high-level overview of various aspects of the disclosure and introduces some of the concepts that are further described in the Detailed Description section below. This summary is not intended to identify key or essential features of the claimed subject matter, nor is it intended to be used in isolation to determine the scope of the claimed subject matter. The subject matter should be understood by reference to appropriate portions of the entirety of this disclosure, any or all drawings and each claim.

[0007] Examples described herein provide for intelligence-driven routing of communication to available workload capacity. Systems and methods for intelligence-driven routing of workload capacity may include storing a workload model in memory regarding a set of different factors associated with user communications. The workload model can then be used to route the workload of agents responding to requests by users for communications. The workload capacity routing can consider different factors to route new requests among agents, and to distribute and monitor capacity of agents engaged with the system responsive to data being received and processed in the system without delays other than communication, processing, or other such resource delays. Each factor may be associated with a measurement of workload. An incoming request for a user communication may be received. Such received request may include information regarding one or more of the factors. A workload measurement may be identified for the requested user communication based on comparing the received request information to the stored workload model. An agent may then be identified as being available to handle the identified workload measurement based on as having a current workload capacity to handle the identified workload measurement. A communication slot defined by the identified workload measurement may be activated for the identified agent. The incoming request may be routed to the identified agent, and a current workload capacity of the identified agent may be dynamically updated based on the identified workload measurement.

[0008] Examples described herein improve the performance of a communication system and devices within the system such as agent devices by efficiently distributing communications among the devices in the system. For example, communication distribution within such a user communication system may not follow normal patterns, and can include communication intensities that are not obvious strictly from the volume of data in a system. Smart analysis of workload can identify patterns of user communication intensity using different factors to predict and identify resource load patterns before capacity issues occur, and route and distribute workload within such a system to avoid capacity issues (e.g., certain agents being overwhelmed or overloaded by a sudden shift in communication intensity while other agents have significant extra capacity).

[0009] Workload models may associate the respective user component and agent component in association with a set of different factors. Each factor may indicative of a measurement of workload defined in a manner to the industry or system type. In a call center, for example, workload may be defined in terms of a standard unit (e.g., call slot) or in terms of percentage. Whereas prior implementations of call center systems may automatically route calls to agents each tasked with handling a predefined number of slots, the present examples can dynamically assign a number of slots to each agent and activate slots for an agent as needed. Such aspects improve the operation of communication systems, networks, and individual devices in the system by optimizing the use of processing and network resources, and therefore improving throughput of the system and avoiding system imbalance and imbalanced use of computing resources which degrade performance.