The overall loss computed by the system, which takes into account DEC loss 36, alignment loss 38, and ranked loss 30, may be represented as follows:

Formula 12:

$$\mathcal{L} = \alpha\mathcal{L}_{\mathtt{DEC}} + \beta\mathcal{L}_{\mathtt{align}} + \gamma\mathcal{L}_{\mathtt{rank}}$$

where *α*, *β*, and *γ* are hyperparameters to trade-off the contribution of different terms.

[0079] The above-described method and stages may be employed for both synthetic and real-world datasets. Three parameters may be imposed for data to be used in the system described herein. The first parameter may include the ability of the data to be formatted as a graph. As long as there exists at least one relationship between either the same type or different types of entities, a graph of the data may be generated. Entities (e.g., users, accounts, merchants, payment devices, etc.) may be formulated as nodes and the relationships may be represented by edges. Edges may be directed or undirected and weighted or unweighted. The graph must be capable of being transformed into adjacency matrices, as described above.

[0080] The second parameter may include the ability of the graphed data to be split into multiple layers. Sets of nodes may be identical across multiple layers and the data may include multiple ways of splitting up a single graph into multiple layers, e.g., by the types of the edges, by the timestamps of the edges, and/or the like. The third parameter may include the existence of anomalies in the data. Because the presently described system is an anomaly-detection system, it is presumed that the data will include meaningful anomalies to be detected. For example, for nodes that represent users, anomalies may include malicious employees inside of organizations, compromised user accounts, fraudulent user activity, and/or the like. For nodes that represent sensors in a network of sensors, an anomaly may include a malfunctioning sensor. It will be appreciated that the type of anomaly may vary and may depend on the system to be analyzed.

[0081] A fourth parameter may be imposed for semi-supervised learning methods, which may require the availability of anomaly feedback. With respect to the availability of anomaly feedback, the dataset may also contain feedback of anomalies, e.g., ground truths of the anomalies. If feedback is not directly available, it may be acceptable that true anomalies can be reasonably derived from the dataset.