

THE INVENTION CLAIMED IS

1. A method comprising:

in response to and immediately after a new entity being added to a fraud detection system:

receiving, with at least one processor, feature data of the new entity and feature data of a plurality of pre-existing entities from at least one data source, wherein the plurality of pre-existing entities are associated with a plurality of pre-existing behavior profiles and the new entity is not associated with a pre-existing behavior profile;

estimating, with at least one processor, an amount of computation time for generating a doppelganger behavior profile for the new entity based on the feature data of the new entity, the feature data of the plurality of pre-existing entities, and a plurality of machine-learning models;

selecting, with at least one processor, a set of parameters of the feature data of the new entity and the feature data of the plurality of pre-existing entities that reduce the amount of computation time for generating the doppelganger behavior profile for the new entity, wherein selecting the set of parameters comprises iteratively performing at least one of a sentiment analysis, a frequency analysis, an accuracy analysis, or any combination thereof on an output of the plurality of machine-learning models to determine an accuracy, a precision, a recall, and/or an F-score for the output and selecting the set of parameters based on the determined accuracy, precision, recall and/or F-score;

generating, with at least one processor, the doppelganger behavior profile for the new entity based at least partially on the set of parameters of the feature data of the new entity and the feature data of the plurality of pre-existing entities, wherein generating the doppelganger behavior profile comprises:

generating a plurality of similarity scores for each entity of the plurality of pre-existing entities based on inputting a portion of the feature data of the new entity and a portion of the feature data of the plurality of pre-existing entities into the plurality of machine-learning models, wherein each similarity score of the plurality of similarity scores defines a similarity between the new entity and an entity of the plurality of pre-existing entities based on an output of a machine-learning model of the plurality of machine-learning models;

determining a single similarity score for each entity of the plurality of pre-existing entities based on the plurality of similarity scores for the entity;

ranking the plurality of pre-existing entities based on the single similarity score for each entity of the plurality of pre-existing entities, to produce a ranking;

selecting two or more pre-existing entities of the plurality of pre-existing based on the ranking;

combining behavior profiles of the two or more preexisting entities to produce the doppelganger behavior profile for the new entity;

assigning, with at least one processor, the doppelganger behavior profile to the new entity; and

monitoring, with at least one processor, activity data of the new entity with respect to the doppelganger behavior profile assigned to the new entity;

analyzing, with at least one processor, the activity data of the new entity with respect to the doppelganger behavior profile assigned to the new entity;

detecting, with at least one processor, in an activity of the new entity at least one of fraud, an anomaly, and malicious intent, based on the doppelganger behavior profile assigned to the new entity; and

automatically suspending, with at least one processor, at least one of an activity of the new entity, access of the new entity to a system, and a function of the new entity in a system in response to detection of the at least one of fraud, an anomaly, and malicious intent.

2. The method of claim 1, further comprising automatically providing, with at least one processor, the new entity access to a pre-existing system based at least partially on the doppelganger behavior profile assigned to the new entity.

3. The method of claim 1, further comprising automatically providing, with at least one processor, a targeted offer to the new entity based at least partially on the doppelganger behavior profile assigned to the new entity.

4. The method of claim 1, further comprising:

converting, with at least one processor, the feature data of the new entity into a feature vector; and

converting, with at least one processor, the feature data of the plurality of pre-existing entities into a feature matrix.

5. The method of claim 1, wherein the new entity is a new account holder, a new enterprise user, or a new enterprise system.

6. The method of claim 1, wherein the feature data comprises at least one of the following parameters: ages of account holders, genders of account holders, credit ratings of account holders, occupations of account holders, account or card type of account holders, applications used by account holders, countries of origin of account holders, locations of account holders, billing addresses of account holders, issuer banks of account holders, issuer bank addresses of account holders, or any combination thereof.

7. The method of claim 1, wherein the feature data comprises at least one of the following parameters: depths of common ancestors in organizational hierarchies of enterprise users, job titles of enterprise users; job descriptions of enterprise users, group memberships of enterprise users, education level of enterprise users, previous employers of enterprise users, connections on social networks of enterprise users, universities of enterprise users, software downloads of enterprise users, browser history of enterprise users, application used by enterprise users, or any combination thereof.

8. The method of claim 1, wherein the feature data comprises at least one of the following parameters: data center locations of enterprise systems, zone placements of enterprise systems, network placements of enterprise systems, application technical contacts of enterprise systems, application

critically ratings of enterprise systems, virtualization technologies used by enterprise systems, software dependencies of enterprise systems, programming languages used by enterprise systems, ports exposed by enterprise systems, underlying operating systems of enterprise systems, communication protocols used by enterprise systems, or any combination thereof.

9. A system comprising at least one server computer including at least one processor, the at least one server computer programmed and/or configured to:

in response to and immediately after a new entity being added to a fraud detection system:

receive feature data of the new entity and feature data of a plurality of pre-existing entities from at least one data source, wherein the plurality of pre-existing entities are associated with a plurality of pre-existing behavior profiles and the new entity is not associated with a pre-existing behavior profile;

estimate an amount of computation time for generating a doppelganger behavior profile for the new entity based on the feature data of the new entity, the feature data of the plurality of pre-existing entities, and a plurality of machine-learning models;

select a set of parameters of the feature data of the new entity and the feature data of the plurality of pre-existing entities that reduce the amount of computation time for generating the doppelganger behavior profile for the new entity, wherein selecting the set of parameters comprises iteratively performing at least one of a sentiment analysis, a frequency analysis, an accuracy analysis, or any combination thereof on an output of the plurality of machine-learning models to determine an accuracy, a precision, a recall, and/or an F-score for the output and selecting the set of parameters based on the determined accuracy, precision, recall and/or F-score;

generate the doppelganger behavior profile for the new entity based at least partially on the set of parameters of the feature data of the new entity and the feature data of the plurality of pre-existing entities, wherein, when generating the doppelganger behavior profile, the at least one server computer is programmed and/or configured to:

generate a plurality of similarity scores for each entity of the plurality of pre-existing entities based on inputting a portion of the feature data of the new entity and a portion of the feature data of the plurality of pre-existing entities into the plurality of machine-learning models, wherein each similarity score of the plurality of similarity scores defines a similarity between the new entity and an entity of the plurality of pre-existing entities based on an output of a machine-learning model of the plurality of machine-learning models;

determine a single similarity score for each entity of the plurality of pre-existing entities based on the plurality of similarity scores for the entity;

rank the plurality of pre-existing entities based on the single similarity score for each entity of the plurality of pre-existing entities, to produce a ranking;

select two or more pre-existing entities of the plurality of pre-existing based on the ranking;

combine behavior profiles of the two or more preexisting entities to produce the doppelganger behavior profile for the new entity;

assign the doppelganger behavior profile to the new entity; and

monitor activity data of the new entity with respect to the doppelganger behavior profile assigned to the new entity;

analyze the activity data of the new entity with respect to the doppelganger behavior profile assigned to the new entity;

detect in an activity of the new entity at least one of fraud, an anomaly, and malicious intent, based on the doppelganger behavior profile assigned to the new entity; and

automatically suspend at least one of an activity of the new entity, access of the new entity to a system, and a function of the new entity in a system in response to detection of the at least one of fraud, an anomaly, and malicious intent.

10. The system of claim 9, wherein the at least one server computer is programmed and/or configured to automatically provide the new entity access to a pre-existing system based at least partially on the doppelganger behavior profile assigned to the new entity.

11. The system of claim 9, wherein the at least one server computer is programmed and/or configured to automatically provide a targeted offer to the new entity based at least partially on the doppelganger behavior profile assigned to the new entity.

12. The system of claim 9, wherein the at least one server computer is programmed and/or configured to:

convert the feature data of the new entity into a feature vector; and

convert the feature data of the plurality of pre-existing entities into a feature matrix.

13. The system of claim 9 wherein the new entity is a new account holder, a new enterprise user, or a new enterprise system.

14. A computer program product comprising at least one non-transitory computer-readable medium including program instructions that, when executed by at least one processor, cause the at least one processor to:

in response to and immediately after a new entity being added to a fraud detection system:

receive feature data of the new entity and feature data of a plurality of pre-existing entities from at least one data source, wherein the plurality of pre-existing entities are associated with a plurality of pre-existing behavior profiles and the new entity is not associated with a pre-existing behavior profile;

estimate an amount of computation time for generating a doppelganger behavior profile for the new entity based on the feature data of the new entity, the feature data of the plurality of pre-existing entities, and a plurality of machine-learning models;

select a set of parameters of the feature data of the new entity and the feature data of the plurality of pre-existing entities that reduce the amount of computation time for generating the doppelganger behavior profile for the new entity, wherein selecting the set of parameters comprises iteratively performing at least one of a sentiment analysis, a frequency analysis, an accuracy analysis, or any combination thereof on an output of the plurality of machine-learning models to determine an accuracy, a precision, a recall, and/or an F-score for the output and selecting the set of parameters based on the determined accuracy, precision, recall and/or F-score;

generate the doppelganger behavior profile for the new entity based at least partially on the set of parameters of the feature data of the new entity and the feature data of the plurality of pre-existing entities, wherein the program instructions that cause the at least one processor to generate the doppelganger behavior profile cause the at least one processor to:

generate a plurality of similarity scores for each entity of the plurality of pre-existing entities based on inputting a portion of the feature data of the new entity and a portion of the feature data of the plurality of pre-existing entities into the plurality of machine-learning models, wherein each similarity score of the plurality of similarity scores defines a similarity between the new entity and an entity of the plurality of pre-existing entities based on an output of a machine-learning model of the plurality of machine-learning models;

determine a single similarity score for each entity of the plurality of pre-existing entities based on the plurality of similarity scores for the entity;

rank the plurality of pre-existing entities based on the single similarity score for each entity of the plurality of pre-existing entities, to produce a ranking;

select two or more pre-existing entities of the plurality of pre-existing based on the ranking;

combine behavior profiles of the two or more preexisting entities to produce the doppelganger behavior profile for the new entity;

assign the doppelganger behavior profile to the new entity; and

monitor activity data of the new entity with respect to the doppelganger behavior profile assigned to the new entity;

analyze the activity data of the new entity with respect to the doppelganger behavior profile assigned to the new entity;

detect in an activity of the new entity at least one of fraud, an anomaly, and malicious intent, based on the doppelganger behavior profile assigned to the new entity; and

automatically suspend at least one of an activity of the new entity, access of the new entity to a system, and a function of the new entity in a system in response to detection of the at least one of fraud, an anomaly, and malicious intent.