manually or through an automated script, regardless of whether their friends are included in the image. Such actions may be suppressed through rate limiting and other policy-based implementation. Implementations of the present disclosure extend these concepts by allowing the social networking system to extract features from the image and probabilistically calculate whether the image is a legitimate natively-captured image (such as one taken from a camera) or a illegitimate non-natively-captured image (such as generated by an application, an advertisement, or a flyer). Other embodiments are readily envisioned by the disclosure and are described in detail below.

[0019] FIG. **1** illustrates an exemplary environment **100** in which a social network may exist and in which objectionable conduct may occur. Environment **100** comprises a server **110**, a network **130**, and user devices **140** and **150**. A social network database **120** is located on the server **110**. The server **110** monitors user activity and determines whether a predetermined policy is violated. The database **120** manages user information that is accessed by users of the social network. The server **110** and the database **120** are discussed further herein. The network **130** provides the medium for communication between the server **110** and user devices **140** and **150**. The network **130** may be the Internet, a wireless or wired network such as a mobile device carrier network, or any other network that can be used for communication between a server and a client.

[0020] In one embodiment, the user device **140** is used by a user initiating the actions being monitored by the server **110**. The user device **150** is used by a recipient of the monitored actions. User devices **140** and **150** may each be any suitable computing device, such as a mobile device, a mobile phone, a cell-phone, a smart-phone, a personal digital assistant, a notebook computer, a laptop computer or a hand-held game console.

[0021] FIG. **2** illustrates one embodiment of the architecture of a server **110** that may be used in a social network. The server **110** comprises the database **120**, a communication module **210**, an accounts module **220**, a policy module **230**, a violation module **240**, and memory **250**. Server **110** also includes a photo store **270** to which users may upload photos to, access photos in, or otherwise interact with photos stored therein via communication module **210**.

[0022] The communication module **210** may be used to communicate with user devices **140** and **150**. The accounts module **220** monitors and/or receives actions from user devices, such as user device **140**. In particular embodiments, accounts module **220** also stores profile information for each particular user. Actions from user device **140** that affect other user devices, such as user device **150**, are time-stamped and the type of action is recorded in memory **250**.

[0023] The policy module **230** stores one or more policies which determine whether a user's actions are objectionable. A policy may be represented by a type of action or a combination of actions and a time dependent function. Such functions may be linear, exponential, polynomial, or step functions, functions of acceleration, i.e., a function of time squared, or any other function that can be used to measure user activity over time. In various embodiments, a representative curve of the policy threshold has an initial steep slope that gradually decreases over time. Particular embodiments of the disclosure pertain to classification or decision functions that map particular inputs with respect to time to a set of actions.

[0024] The policy may have different thresholds of violation for different degrees of the same type of action, e.g., a warning threshold which results in some warning to the offending user, and a suspension threshold which removes rights from the user. A suspension may be temporary, for example for hours, days, or months, or may require a system operator to reactivate a user's account prior to further actions. A suspension may help curtail spammers from abusing the social network. In various embodiments, the time of a temporary suspension is extended if a user attempts to repeat the type of action that resulted in suspension of the account.

[0025] The violation module **240** determines whether a user has violated the policy. The violation module **240** compares the number and type of actions recorded by the accounts module **220** with the policy stored in the policy module **230**. Upon reaching a threshold, the violation module **240** may respond by requesting the communication module **210** to issue a warning to the user and/or by altering a user record in the database **120** to suspend the user's account.

[0026] The server **110** may include memory **250** to store applications or modules and other information. Photo spam detector **260** residing on server **110** includes instructions embodied in software code operable to detect illegitimate photos and illegitimate photo-tagging operations, and suppress the delivery of such photos or tags. Photo spam detector **260** is operably connected to photo store **270**, and may access information stored in accounts module **220** and policy module **230**. In some embodiments, applications, modules, and/or database **120** may reside in memory on several servers (not shown).

[0027] FIG. **3** illustrates exemplary individual software modules within photo spam detector **260**. Image feature extraction module **301** applies image processing algorithms or computer vision algorithms to extract features from a given image. Such features may include color distributions, histograms for one or more colors or gamma levels, the aspect ratio of the photo image, the compression ratio of the image, edge detection, texture detection, orientation, and the like. Methods of image feature extraction are well-known in the art and include but are not limited to frequency domain filtering, wavelet transforms, and converting the image to vector form (such as SVG). This disclosure contemplates any suitable method of image feature extraction.

[0028] The various features extracted by image feature extraction module **301** may give photo spam detector **260** insight as to whether the image in question is a legitimate natively-captured image or an illegitimate non-natively captured image. In particular embodiments, color histograms may provide useful insights. Images with large regions of unnatural colors, such as pure black or pure white, are less likely to be natively-captured images. These images, in particular embodiments, are much more likely to be computer-generated graphics like the example image of FIG. **4**B. In particular embodiments, the color distribution lends insight into the legitimacy of the image. For example, even where a clever spammer may utilize a color gradient as a background, the linearity of the gradient along any given axis or distance may indicate that the image is illegitimate. Images with linear gradients are much less likely to be captured natively. In particular embodiments, edge detection may be utilized to score the legitimacy of the image. For example, natural images do not have clear, well defined