



(51) International Patent Classification:

H04L 51/02 (2022.01) *G06F 40/35* (2020.01)
G06F 40/30 (2020.01) *G06F 40/56* (2020.01)

(21) International Application Number:

PCT/EP2024/063640

(22) International Filing Date:

17 May 2024 (17.05.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

18/320,250 19 May 2023 (19.05.2023) US

(71) Applicant: **BITDEFENDER IPR MANAGEMENT LTD** [CY/CY]; 59-61 Acropolis ave. 3rd floor, office 302, Nicosia, 2012 (CY).

(72) Inventors: **BRAD, Florin M.**; Str. Petre Liciu, nr. 4, Jud. Prahova, 105600 Câmpina (RO). **MANOLACHE, Andrei M.**; Mitropolit Veniamin Costache 1, Jud. Neamț, 615200 Târgu-Neamț (RO). **BURCEANU, Elena**; Str. Constantin Rădulescu-Motru nr. 4, bl.1, sc. 2, et. 6, ap. 58 Sector 6, 040361 București (RO).

(74) Agent: **TULUCA, Doina**; Bd. Lacul Tei nr. 56, bl 19, sc B, ap 52, sector 2, 020392 București (RO).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,

(54) Title: SYSTEMS AND METHODS OF DETECTING CHATBOTS

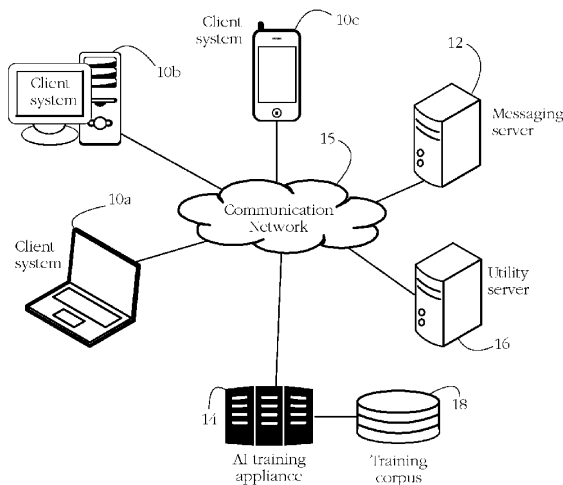


FIG. 1

(57) Abstract: Some embodiments determine whether an entity engaging in online conversations comprises a chatbot. A challenge message is constructed according to a current context of an ongoing conversation, by employing a generative language model to determine a plausible conversation sequel and subsequently distorting the respective sequel. The challenge is added as a new message to the respective conversation. The chatbot detector may then determine whether a conversation partner is a chatbot according to a similarity between a response to the challenge received from the respective conversation partner and an artificially generated, plausible response to the undistorted conversation sequel. Distortions applied to construct the challenge are deliberately crafted to send a chatbot on a markedly different trajectory.

[Continued on next page]

TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))