A NEW ANALYSIS OF CODE RED AND WITTY WORMS BEHAVIOR

NIKOLAY KYURKCHIEV¹, ANTON ILIEV², ASEN RAHNEV³, AND TODORKA TERZIEVA⁴

¹,²,³,⁴Faculty of Mathematics and Informatics
University of Plovdiv Paisii Hilendarski
24, Tzar Asen Str., 4000 Plovdiv, BULGARIA

ABSTRACT: In this note we give how can be make a new precise analysis of the situation with spreading of Code Red worm as well Witty worm. These worms were actively spread from 00:00 UTC July 19, 2001 to 00:00 UTC July 20, 2001 and on March 19, 2004, at approximately 8:45 p.m. Pacific Standard Time (PST) respectively. Here we give a new way of treating these epidemics using Dagum-II sigmoid function [1]–[7].

AMS Subject Classification: 97N50
Key Words: Dagum-II cumulative distribution function (DIICDF), shifted Heaviside function, Hausdorff distance, Code Red worm, Witty worm, upper and lower bounds

Received: November 6, 2018; Accepted: January 11, 2019; Published: January 13, 2019; doi: 10.12732/caa.v23i2.3

1. INTRODUCTION

For approximation and modelling of such data in the literature has become necessary and used: a finite population the classical simple epidemic model
\[ \frac{dJ(t)}{dt} = \beta J(t)[N - J(t)], \]

where \( J(t) \) is the number of infected hosts at time \( t \), \( N \) is the size of population, \( \beta \) is the infection rate. Initially, \( t = 0, J(0) \) hosts are infected and the other \( N - J(0) \) hosts are all susceptible.

With using epidemic models in many cases cannot be received expected good fitting.

Other models that can be used are given in [17].

For instance (see, [17]), let \( I(t) \) denote the number of infectious at time \( t \). We use \( R(t) \) to denote the number of removed hosts from previously infectious at time \( t \).

Let \( J(t) \) denote the number of infected hosts by time \( t \), no matter whether they are still in infectious state or have been removed, i.e. \( J(t) = I(t) + R(t) \).

The Kermack–Mckendrick general epidemic model is:

\[
\frac{dJ(t)}{dt} = \beta J(t)[N - J(t)] \\
\frac{dR(t)}{dt} = \gamma I(t) \\
J(t) = I(t) + R(t) = N - S(t),
\]

where \( \beta \) is the infection rate; \( \gamma \) is the rate of removal of infectious hosts from circulation; \( S(t) \) is the number of susceptible hosts at time \( t \); \( N \) is the size of population.

The so-called ”two-factor worm model” gives better results [17], but numerical analysis is very difficult due to the presence of many indeterminate dynamic parameters.

The number of sigmoid functions which are used for data analysis in domain of population dynamics, biostatistics, debugging and test theory are explored in the following monographs [10]–[14]. Approximation of step function with cumulative function of type Dagum-II for wealth data is studied in details [15].

We will illustrate the advances of this differential model for approximation and modelling of data (see Fig. 3) for Code Red worm [16], [17] as well as for Witty worm for entire world and for USA [33].
Definition 1. The (basic) step function is:

\[
\tilde{h}_{r,t_0}(t) = \begin{cases} 
    r, & \text{if } t < t_0, \\
    [r,1], & \text{if } t = t_0, \\
    1, & \text{if } t > t_0,
\end{cases}
\]  

(1)

usually known as shifted Heaviside function.

Definition 2. [8], [9] The Hausdorff distance (the H–distance) \( \rho(f,g) \) between two interval functions \( f, g \) on \( \Omega \subseteq \mathbb{R} \), is the distance between their completed graphs \( F(f) \) and \( F(g) \) considered as closed subsets of \( \Omega \times \mathbb{R} \).

More precisely,

\[
\rho(f,g) = \max \left\{ \sup_{A \in F(f)} \inf_{B \in F(g)} ||A - B||, \sup_{B \in F(g)} \inf_{A \in F(f)} ||A - B|| \right\},
\]

(2)

wherein \( ||.|| \) is any norm in \( \mathbb{R}^2 \), e. g. the maximum norm \( ||(t,x)|| = \max\{|t|,|x|\} \); hence the distance between the points \( A = (t_A,x_A), B = (t_B,x_B) \) in \( \mathbb{R}^2 \) is \( ||A - B|| = \max(|t_A - t_B|,|x_A - x_B|) \).

So far as differential model connected to cumulative function of type Dagum-II is described in the literature [1] here we will point the solution of differential equation

\[
F(t) = \left( 1 + \left( \frac{t}{b} \right)^{-a} \right)^{-p}.
\]

(3)

This approach was further developed in a series of papers on generating systems for income distribution [3]–[6].

For other results, see [7].

The extended Dagum-II type distribution has the following cumulative function(D–II CDF):

\[
F(t) = r + (1 - r) \left( 1 + \left( \frac{t}{b} \right)^{-a} \right)^{-p}; \quad t > 0,
\]

(4)

where \( a, b, p > 0 \) and \( r \in (0,1) \).

When modelling such processes it is important to study the characteristic "supersaturation".
Let \( t_0 = b \left( \left( \frac{1}{2} \right)^{-\frac{1}{p}} - 1 \right)^{-\frac{1}{a}} \), then \( F(t_0) = \frac{1+r}{2} \). We will use the following theorem for Dagum-II cumulative distribution function:

**Theorem 1.** [15] Let

\[
\alpha = -\frac{1-r}{2}; \quad \beta = 1 + \frac{ap(1-r)}{b} \left( \frac{1}{2} \right)^{1+p} \left( \left( \frac{1}{2} \right)^{-\frac{1}{p}} - 1 \right)^{\frac{1+a}{a}}.
\]

The \( H \)-distance \( d \) between the function \( \tilde{h}_{r,t_0} \) and the function \( F^* \) can be expressed in terms of the parameters \( a, b, p, r \) for any real \( \beta \geq e^{\frac{1.05}{2.1}} (1-r) \) as follows:

\[
d_l = \frac{1}{2.1 \frac{\beta}{1-r}} < d < \frac{\ln(2.1 \frac{\beta}{1-r})}{2.1 \frac{\beta}{1-r}} = d_r.
\]  

\[ (6) \]

2. MAIN RESULTS

2.1. ANALYSIS OF CODE RED WORM INFECTION BEHAVIOR

Here we will give an application of cumulative sigmoid when provide analysis of this real data [16], [17].

From data analysis it follows that it is worthy to use the model:

\[
\tilde{F}(t) = \omega \left[ r + (1-r) \left( 1 + \left( \frac{t}{b} \right)^{-a} \right)^{-p} \right].
\]

When solving this task in programming environment Mathematica we receive the results (see, Fig. 1 and Fig. 2):

\[ \omega = 341015; \quad r = 0.00773133; \quad a = 10.8297; \quad b = 484.727; \quad p = 1274.96. \]

It can be seen that model function that we choose gives good fitting. This give basis for specialists working in this area for analyzing similar functions with appropriate modifications of cumulative function of type Dagum-II.

We will use the following data which we photographed for the situation that was happen on July 19, 2001, [16] where the first coordinate is time with step 5 minutes and second coordinate is number of infected hosts:
CODE RED AND WITTY WORMS BEHAVIOR

\{665, 4679\}, \{670, 4718\}, \{675, 4767\}, \{680, 4797\}, \{685, 4830\}, \{690, 4873\}, \{695, 4920\}, \{700, 4962\}, \{705, 4996\}, \{710, 5040\}, \{715, 5083\}, \{720, 5140\}, \{725, 5202\}, \{730, 5253\}, \{735, 5310\}, \{740, 5378\}, \{745, 5456\}, \{750, 5525\}, \{755, 5614\}, \{760, 5720\}, \{765, 5840\}, \{770, 5989\}, \{775, 6155\}, \{780, 6352\}, \{785, 6582\}, \{790, 6864\}, \{795, 7231\}, \{800, 7637\}, \{805, 8142\}, \{810, 8716\}, \{815, 9412\}, \{820, 10229\}, \{825, 11186\}, \{830, 12273\}, \{835, 13502\}, \{840, 14950\}, \{845, 16688\}, \{850, 18632\}, \{855, 20842\}, \{860, 23362\}, \{865, 26267\}, \{870, 29500\}, \{875, 33084\}, \{880, 37207\}, \{885, 41834\}, \{890, 46966\}, \{895, 53012\}, \{900, 59782\}, \{905, 67109\}, \{910, 74905\}, \{915, 83125\}, \{920, 91814\}, \{925, 100917\}, \{930, 110484\}, \{935, 120697\}, \{940, 131103\}, \{945, 141400\}, \{950, 151896\}, \{955, 162369\}, \{960, 172670\}, \{965, 182971\}, \{970, 193003\}, \{975, 202851\}, \{980, 208254\}, \{985, 209439\}, \{990, 211154\}, \{995, 211162\}, \{1000, 222270\}, \{1005, 222301\}, \{1010, 226442\}, \{1015, 231915\}, \{1020, 231963\}, \{1025, 231970\}, \{1030, 231971\}, \{1035, 235749\}, \{1040, 243244\}, \{1045, 249363\}, \{1050, 254182\}, \{1055, 258904\}, \{1060, 262950\}, \{1065, 266407\}, \{1070, 270566\}, \{1075, 273155\}, \{1080, 275551\}, \{1085, 277719\}, \{1090, 279974\}, \{1095, 281542\}, \{1100, 283066\}, \{1105, 285319\}, \{1110, 286755\}, \{1115, 288206\}, \{1120, 289618\}, \{1125, 290878\}, \{1130, 292924\}, \{1135, 294269\}, \{1140, 295579\}, \{1145, 296702\}, \{1150, 297971\}, \{1155, 298875\}, \{1160, 300046\}, \{1165, 301232\}, \{1170, 302573\}, \{1175, 303922\}, \{1180, 305103\}, \{1185, 306286\}, \{1190, 307222\}, \{1195, 308014\}, \{1200, 308705\}, \{1205, 309787\}, \{1210, 310683\}, \{1215,
Figure 1: Epidemic data of Code Red worm fitted with Dagum-II cumulative distribution function

\[
\begin{align*}
312041, & \{1220, 312907\}, \{1225, 313707\}, \{1230, 314729\}, \{1235, 315956\}, \{1240, 317079\}, \{1245, 317914\}, \{1250, 318610\}, \{1255, 319454\}, \{1260, 320375\}, \{1265, 321177\}, \{1270, 321993\}, \{1275, 322706\}, \{1280, 323644\}, \{1285, 324601\}, \{1290, 325398\}, \{1295, 326217\}, \{1300, 327050\}, \{1305, 327676\}, \{1310, 328422\}, \{1315, 328990\}, \{1320, 329833\}, \{1325, 330892\}, \{1330, 331442\}, \{1335, 332030\}, \{1340, 332952\}, \{1345, 333672\}, \{1350, 334172\}, \{1355, 334843\}, \{1360, 335627\}, \{1365, 336324\}, \{1370, 337002\}, \{1375, 337606\}, \{1380, 338427\}, \{1385, 339866\}, \{1390, 340628\}, \{1395, 341015\}, \{1400, 341015\}, \{1405, 341015\}, \{1410, 341015\}, \{1415, 341015\}, \{1420, 341015\}, \{1435, 341015\}.
\end{align*}
\]
Figure 2: Dagum-II cumulative distribution function for epidemic data of Code Red

Figure 3: Global epidemic of Code Red worm distribution
2.2. ANALYSIS OF WITTY WORM INFECTION BEHAVIOR

Here we will give an application of cumulative sigmoid when provide analysis of this real data [33], see Fig. 4.

From data analysis it follows that it is worthy to use Dagum-II model (7). When solving this task in programming environment Mathematica we receive the results (see, Fig. 5 and Fig. 7 respectively for entire World and USA spreading of Witty worm):

- for entire World spreading parameters are

\[
\omega = 12372; a = 2.18621; b = 34.962; p = 0.610061; r = 0.00435232,
\]

and

- for USA spreading parameters are
\[ \omega = 6063; \quad a = 2.24077; \quad b = 31.8307; \quad p = 0.606965; \quad r = 0.019863. \]

For USA spreading the predictive power (PP) criterion:

\[ PP = \sum \left( \frac{\hat{F}(t_i) - y_i}{y_i} \right)^2 \]

we find \( PP = 0.0487828 \).

Again it can be seen that model function that we choose gives good fitting. So similar functions with appropriate modifications of cumulative function of type Dagum-II will be of interest of specialists working in this area.

We will use the following data which we photographed for the situation that was happen on March 19, 2004, at approximately 8:45 p.m. Pacific Standard Time (PST), [33] where the first coordinate is time with step 5 minutes and second coordinate is number of infected hosts (see Fig. 6 and Fig. 8):

\[
\text{data}_{\text{World}} = \{(0.1,150),\{5,5869\},\{10,2141\},\{15,3637\},\{20,5312\},
\{26,6602\},\{31,7562\},\{36,8340\},\{41,8941\},\{46,9389\},\{51,9734\},
\{56,10060\},\{61,10349\},\{66,10586\},\{71,10800\},\{76,11169\},
\{86,11362\},\{91,11532\},\{96,11684\},\{101,11823\},\{106,11972\},
\{111,12118\},\{116,12256\},\{121,12372\}\}
\]

and

\[
\text{data}_{\text{USA}} = \{(0.1,150),\{5,576\},\{10,1236\},\{15,1963\},\{20,2973\},
\{26,3488\},\{31,3953\},\{36,4343\},\{41,4630\},\{46,4825\},\{51,4986\},
\{56,5153\},\{61,5280\},\{66,5380\},\{71,5468\},\{76,5590\},\{86,5706\},
\{91,5769\},\{96,5831\},\{101,5877\},\{106,5939\},\{111,5989\},
\{116,6033\},\{121,6063\}\}.
\]
Figure 5: Epidemic data for entire world of Witty worm fitted with Dagum-II cumulative distribution function

Figure 6: Global epidemic of Witty worm for entire world
Figure 7: Epidemic data for USA of Witty worm fitted with Dagum-II cumulative distribution function

Figure 8: Global epidemic of Witty worm for USA
Conclusions

The reader can see the advantages of suggested by us using of Dagum-II cumulative distribution function. These results will be of interest for specialists in this modern scientific branch [16]–[62].

ACKNOWLEDGMENTS

This work has been supported by D01-205/23.11.2018 National Scientific Program “Information and Communication Technologies for a Single Digital Market in Science, Education and Security (ICTinSES)”, financed by the Ministry of Education and Science, Bulgaria.

REFERENCES


