










Analysis of the Number of Sides of Voronoi Polygons in PassPoint

Lisset Suárez-Plasencia¹ , Joaquín A. Herrera-Macías² ,
Carlos M. Legón-Pérez³ , Raisa Socorro-LLanes⁴ , Omar Rojas⁵ ,
and Guillermo Sosa-Gómez⁵  

¹ Departamento de Matemática-Física, Universidad de Artemisa, Artemisa, Cuba
l.suarez2@uart.edu.cu

² Departamento de Matemática, Universidad de Matanzas, Matanzas, Cuba
joaquin.herrera@umcc.cu

³ Facultad de Matemática y Computación, Instituto de Criptografía,
Universidad de la Habana, Habana, Cuba
clegon58@gmail.com

⁴ Facultad de Informática, Universidad Tecnológica de la Habana, Habana, Cuba
raisa@ceis.cujae.edu.cu

⁵ Facultad de Ciencias Económicas y Empresariales, Universidad Panamericana,
Álvaro del Portillo 49, 45010 Zapopan, Jalisco, Mexico
{orojas,gsosag}@up.edu.mx

Abstract. The probabilistic distribution of the characteristics of Voronoi polygons has been extensively studied due to its many areas of application. In various works that differ in the number of polygons generated and in the size of their regions, it is concluded that the expected value of the characteristic number of sides of Voronoi polygons is equal to 6. In this work, this characteristic in the polygons generated by the graphical passwords of the graphical authentication system *PassPoint* is studied. Its distribution is estimated and it is shown that the expected value of the number of sides of the Voronoi polygons in this scenario differs from previous works. The effectiveness of this feature is evaluated to detect weak graphical passwords made up of grouped dots. They are to be detected by estimating the entropy of the number of sides and by the expected value of the number of sides. It is concluded that the distribution of the number of sides in this scenario does not the 3-parameter gamma distribution reported in previous work or any of 61 distributions that were tested, and that the entropy and the expected value of the number of sides are not efficient for the detection of weak graphical passwords of *PassPoint* formed by 5 grouped points.

Keywords: Passpoint · Voronoi polygons · Entropy

1 Introduction

The antecedents of the Voronoi diagrams date back to the 17th century, when the French mathematician Rene Descartes, in his work *Principia Philosophiae*

published in 1644, describes a partition of the universe's discs into 'vortices'. Although he does not explicitly define his vortexes in the manner of Voronoi cells (regions), his work is conceptually very similar [2,36]. The first references on the subject were known in 2003, because his idea did not coincide with that of the mathematicians of that time. In the 19th century, the Voronoi diagram was conceptually formalized by the German mathematician Johann Peter Gustav Lejeune Dirichlet for the 2D and 3D cases, and extended at the beginning of the next century by the Russian mathematician Georges Voronoi for the n -dimensional case [36]. These diagrams are nothing more than a geometric structure studied in computational geometry, which represents approximate information about a set of points called sites or generators. Voronoi (or Thiessen) diagrams or polygons (as they are also often called) are known in different ways due to their immense applications in various branches of science, including computation, meteorology, physics, geology, crystallography, anthropology, among others [2,36]. Recently, in [26] and [39] they have been used in graphical authentication and robotics respectively.

The dual of a 2D Voronoi diagram is a Delaunay triangulation (or Delaunay mosaic), and was exposed before 1872 by the French mathematician Charles Delaunay [36]. These diagrams and their dual, present a set of characteristics determined by the properties of randomness or dependence of the initial set of points. Polygons where their points follow a random distribution are called *Voronoi Poisson polygons* [8–11,16,17,20].

One of the main applications of Voronoi polygons using their characteristics is to evaluate randomness or detect the existence of patterns in the initial set of points [37]. According to on the behavior of the points distributed in the plane, the spatial point patterns are classified as random (homogeneous Poisson point process), regular (uniform or an inhibiting pattern) or grouped (aggregated), see [3,11,18,32,37,47].

In [11], a set of these characteristics was investigated for the existence of patterns, however the exclusion of the number of sides of the Voronoi polygons from the set of investigated characteristics is not argued. On the other hand, in [22] it was shown that the K-Ripley function and the distance to the nearest neighbor, two of the most used tests in spatial randomness, are inefficient to detect clustering in graphical passwords in the system *PassPoint* (which is based on the user remembering 5 dot patterns on an image selected as their password) [29,41,48]. In this system, a password is considered weak if the 5 points selected by the user do not follow a random distribution. The main types of non-randomness that may be present between the points in that case are: grouping, regularity, smoothness, and symmetry. By the results of [11] and [22] the efficiency of the characteristic number of sides of Voronoi polygons, to detect graphical passwords with points grouped in *PassPoint* is investigated in this work.

The numerous applications of the Voronoi polygons in different areas of knowledge have generated a great variety of studies on their characteristics and their proballistic distribution. These distributions, as well as those of the characteristics of the Delaunay triangulations in the two-dimensional case, are unknown

in many cases, and it is necessary to apply simulation techniques to estimate them [7, 8, 11, 13, 14, 23, 24, 28, 38, 46].

In [10, 20, 21], some theoretical results associated with the distribution of the number of sides of the Voronoi polygons. In [7, 8, 11, 13, 14, 23, 24, 28, 46] the distribution of the number of sides of the Voronoi Poisson polygons is estimated by simulation. Despite being generated in these studies between 200 and 208,969,210 of Voronoi Poisson polygons in a given region of the plane, the expected value does not vary, it is always 6 [7, 8, 11, 13, 14, 23, 24, 28, 46]. In [6], to measure the level of uniformity of the obtained polygons, the probability P_N is defined as the proportion of polygons with N sides and to this distribution $\{P_N | \sum_N P_N = 1\}$ of the number of sides, the Voronoi entropy $S_{vor} = H(P_N) = -\sum P_N \log P_N$ is calculated, which allows quantify the ordering of the set of points in the plane or the cells around these points.

On the other hand, in graphical authentication, entropy is used as a measure of complexity, the amount of information in an image, or the security of a password. This is especially useful when estimating whether an image is useful for the authentication process using the *PassPoint* system. In [27], it is concluded that image passwords with high entropy are easy to forget, or what is the same difficult to remember the password.

In this work a state of the art distribution of the characteristics of Voronoi polygons is presented, in particular on the characteristic number of sides. It is investigated, for the first time as far as we know, in a very peculiar scenario: the polygons generated by the graphical passwords used in the *PassPoint* graphical authentication system, where it is only possible to generate 5 polygons on a rectangular area of the flat. The distribution of this characteristic in this scenario is estimated and its effectiveness in detecting weak graphical passwords formed by grouped points is evaluated.

The work is structured in 6 sections: Section 1 shows the Introduction; Section 2 is composed by Voronoi polygons and *PassPoint*; Section 3 shows the background of the distribution of some characteristics of the Voronoi polygons and specifically the background of the distribution of the number of sides of the Voronoi Poisson polygons; Section 4 presents our main contribution: Analysis of the number of sides of Voronoi polygons in *PassPoint*; Section 5 show the comparison with previous works; and finally in Section 6 the conclusions are presented.

2 Preliminaries

2.1 Voronoi Polygons

Voronoi diagrams are a geometric construct that, given a set $P = \{p_1, p_2, \dots, p_n\}$ of n points, called sites, allows to build a partition of the Euclidean plane in a set of n disjoint regions, so that each region $V(p_i)$ corresponds to a single site p_i . The points q belonging to a given region $V(p_i)$ fulfill the property of being at a lower (Euclidean) distance from the site p_i corresponding to that region than to any other site p_j ; i.e., $d(q, p_i) \leq d(q, p_j)$, $\forall p_i \neq p_j, 1 \leq i, j \leq n$. In Voronoi diagrams not all regions are bounded, the bounded ones are known as

closed Voronoi polygons (V.P.) or convex polygons, and the unbounded ones as unbounded regions or open Voronoi polygons. The boundaries of the Voronoi regions are defined by bisectors joining each pair of sites, and the point of intersection between the bisectors is called the Voronoi vertex. In [5, 15, 19, 30, 38, 45], the Voronoi regions are also often referred to as Voronoi cells, the boundaries as edges, the sites by generating points, and the set formed by these points is called the generator set.

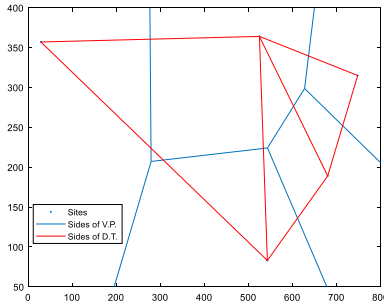


Fig. 1. Representation of a Voronoi diagram and its dual.

The dual of a two-dimensional Voronoi diagram is a Delaunay triangulation (D.T.), see Fig. 1. This triangulation is performed by connecting to the closest vertices, satisfying that all circumscribed circles in the network of triangles are empty, this restriction is known as the Delaunay condition. If P is a randomly generated set of the plane, then the Voronoi polygons and Delaunay triangulations are random, called Voronoi polygons and Delaunay Poisson triangulations [7, 8, 11, 13, 14, 23, 24, 28, 38, 46].

Voronoi polygons and their Delaunay triangulations have a set of characteristics determined by the properties of randomness or dependence on the initial set of points. In the two-dimensional case, these characteristics are the following [11, 38]:

- N number of sides (edges or vertices) of the polygons
- Length L_V of a side of a Voronoi polygon
- Length L_D of a side of a Delaunay triangle
- Distance R between a site and a vertex of its Voronoi polygon (R radius of a circle circumscribed in a Delaunay triangle)
- Area A_D and perimeter P_D of a Delaunay triangle
- Area A_V and perimeter P_V of a Voronoi polygon
- Interior angle α_{int} of a Delaunay triangle
- Minimum angle α_{min} , mean angle α_{med} and maximum angle α_{max} of a Delaunay triangle

In [11] the characteristics were analyzed, α_{max} , α_{med} , α_{int} , L_V , R , A_D , P_D , L_D and α_{min} , given an initial set of 100 clustered or regular patterns in a square

unit, concluding that the first eight characteristics are more competent to detect clustering and the last regularity [11].

2.2 PassPoint

The *PassPoint* [29, 41, 48] system is one of the most advantageous techniques of the *cued-recall* type in graphical authentication, due to its security and usability. This technique requires the user to select an ordered set of 5 pixels in an image as their password in the registration phase, and in the authentication phase they must select “approximately” the same pixels, and in the same order in which they were selected in the registration phase.

In graphical authentication with *PassPoint*, if the set of points (pixels) selected by the user as their graphical password do not follow a random pattern, then said graphical password is considered weak, as it can be compromised by so-called attacks from dictionaries [40, 41]. For this reason, it is necessary to detect user selection of weak passwords. For this reason, it would be useful to evaluate the effectiveness of the Voronoi polygon characteristics in detecting patterns in *PassPoint*.

3 Background of the Distribution of Some Characteristics of the Voronoi Polygons

For many features of Voronoi polygons and Delaunay Poisson triangulations, their distributions are unknown and have been approximated by simulation, but in some cases theoretical results are already known. In [16] and [35], they report the estimated probability density functions for the minimum angle α_{min} and the edge length L_D of a Delaunay Poisson triangle with intensity $\lambda = \frac{N(A)}{|A|}$, ($N(A)$ number of points distributed in the study area A) [18] respectively. In [23] and [46], they obtained the first four moments of N , A_V , P_V and α_{int} from a Voronoi polygon and later adapted their histograms to a Gamma generalized three-parameter distribution; to estimate the parameters of the generalized Gamma distribution of three parameters in [46] they used the maximum likelihood estimator. A summary of various previous works, before the year 2000, on the study of some of the characteristics of Voronoi Poisson polygons by simulation is found in Table 5.4.1 of [38]. For the number of sides, Hayen and Quine in [20, 21] presented an integral formula for \hat{p}_3 (probability associated with the number of side 3), obtaining in [20] a value of 7 decimal places and soon in [21] the improved value for 10 decimal places, $\hat{p}_3 = 0.0112400129$. In 2002, Calka [10] using a technique based on the famous formula given by Slivnyak in [34], proved an explicit expression (an integral formula) for the distribution function of the number of sides of a polygon of Voronoi Poisson, where the value of \hat{p}_3 in [20, 21] matches the value of [10].

3.1 Background of the Distribution of the Number of Sides of the Voronoi Poisson Polygons

In previous works that estimate the distribution of the number of sides of Voronoi Poisson polygons through simulation, two fundamental variants of estimating the number of sides of Voronoi polygons are observed:

- The first variant consists of the simulation of n polygons in n iterations, generating in each of the iterations a set of m random points in a given rectangular region, in which only the number of sides of the polygon is extracted Voronoi associated to the point closest to the center of the region to avoid the edge effect.
- In the second variant, n polygons are generated in a set of the plane, with n relatively “big” neglecting the polygons that are partially bounded by an edge of the study area, calling said edge cause effect [33].

Despite their different estimation methods, their probabilities roughly correspond to the theoretical distribution of the number of sides of the Voronoi polygons found in [10].

Variant 1. In [14], Crain generated a total of 46,000 Voronoi Poisson polygons, to which he added the results of the generation of 11,000 polygons that he had previously published in [13], conducting a study of the number of sides of the 57,000 Voronoi Poisson polygons associated with the points closest to the center. For the simulation of these polygons, he generated a set of 35 random points in a square unit because it was the maximum number of points that could be generated when compiling. Hinde and Miles in [23], to estimate with better precision the properties of the distribution of the number of sides, simulated in n iterations, 2,000,000 Voronoi Poisson polygons associated to the points closest to the center of the rectangle, in each one of them they generated over a rectangular region of the plane Voronoi polygons with intensity $\lambda = 100$. In [24], they simulated 100 Voronoi Poisson polygons in a square of dimensions 25×25 , with units of arbitrary length and intensity $\lambda = 0.16$. To do this, they subdivided the initial square into squares of 5×5 , and in each one they generated 4 sites such that the minimum distance to any other site of the square of 25×25 was greater than a set value, in the case of random points they established the parameter from $\delta = 0.0$ up to $\delta = 0.1$. Kumar and Kurtz [28] reported a simulation of 650,000 Voronoi Poisson polygons in 650,000 iterations, for this they defined a square region in which they generated 100 random points with one of the points in the center of the square, then they calculated the properties of the Voronoi polygons associated with the center point. Tanemura [46] performed basically the same procedure as Hinde and Miles [23], but unlike them, he generated 10,000,000 Voronoi Poisson polygons, for an intensity $\lambda = 200$, in a given region of the plane simulated a number of Voronoi Poisson polygons by estimating only the number of sides of the polygons corresponding to the center point.

Variant 2. In [8], the antecedent with the highest number of simulated polygons is reported, Brakke reports a simulation of 208,969,210 Voronoi Poisson polygons in the plane. In a square unit he generated 88 polygons per second, for a total of approximately 3,801,600 polygons per day, simulating the number of polygons estimated in 55 days. Schmid and Leitner in [44], simulated 100,000 in 100,000 Voronoi Poisson polygons in a rectangular area of approximately 33×17 , sampling up to 1,000,000 Voronoi polygons. To estimate the number of sides of the observed Voronoi polygons, the edges of the study area were not taken into account, the open polygons being discarded by the edge effect. In [7], Bormashenko, Legchenkova, and Frenkel generated a set of 200 random points in a circle of a given diameter, obtaining a Voronoi entropy equal to 1.65.

The studies carried out in [7, 8] and [44] of the number of sides of the Voronoi Poisson polygons, unlike the rest of the antecedents, generate in the same iteration a relatively “large quantity” of polygons. However, their probabilities correspond approximately to the previous ones and the expected value associated with this characteristic is always 6, (see Table 1).

Table 1. Distribution and expected value of the number of sides the Voronoi polygons of the antecedents, except those obtained in [7, 44] because they do not explicitly give their exact values, and those of [24] are not known.

| N | \hat{p}_N (Crain, 1978) | \hat{p}_N (Hinde & Miles, 1980) | \hat{p}_N (Kumar & Kurtz, 1993) | \hat{p}_N (Calka, 2002) | \hat{p}_N (Tanemura, 2003) | \hat{p}_N (Brakke, 2005) |
|--------|---------------------------|-----------------------------------|-----------------------------------|---------------------------|------------------------------|----------------------------|
| 3 | 0.011000 | 0.01131 | 0.01100 | 0.011240 | 0.01125 | 0.01125 |
| 4 | 0.107800 | 0.10710 | 0.10710 | 0.106838 | 0.10685 | 0.10683 |
| 5 | 0.259400 | 0.25910 | 0.26000 | 0.259460 | 0.25941 | 0.25945 |
| 6 | 0.295200 | 0.29440 | 0.29400 | 0.294730 | 0.29479 | 0.29471 |
| 7 | 0.198400 | 0.19910 | 0.19900 | 0.198770 | 0.19884 | 0.19880 |
| 8 | 0.089600 | 0.09020 | 0.09000 | 0.089700 | 0.09003 | 0.09012 |
| 9 | 0.029600 | 0.02950 | 0.03000 | 0.029500 | 0.02963 | 0.02964 |
| 10 | 0.007510 | 0.00743 | 0.00700 | 0.000000 | 0.00743 | 0.00745 |
| 11 | 0.001420 | 0.00149 | 0.00150 | 0.000000 | 0.00149 | 0.00148 |
| 12 | 0.000175 | 0.00025 | 0.00023 | 0.000000 | 0.00025 | 0.00024 |
| 13 | 0.000053 | 0.00003 | 0.00004 | 0.000000 | 0.00003 | 0.00003 |
| $E[N]$ | 6 | 6 | 6 | 6 | 6 | 6 |

Regarding its distribution, as early as in 1980, Hinde and Miles shown that the distribution of the number of sides adjusts to a generalized three parameter Gamma (3P) distribution. Later, in 2003, Tanemura [46] estimated the parameters of said distribution, where $\hat{a} = 0.96853$, $\hat{b} = 3.80078$ and $\hat{c} = 20.86016$.

4 Analysis of the Number of Sides of Voronoi Polygons in *PassPoint*

In the aforementioned previous works, the different scenarios of each study are reflected to estimate the number of sides of the Voronoi polygons, avoiding the edge effect [33] in each one. In [25] they analyzed the relationship between the proportion of points that must be excluded and the number of points in a given area using Monte-Carlo simulation, turning out to be very few points not excluded if the number of points is less than 25.

4.1 Differences of the Polygons of This Scenario with Previous Studies

In this work, the study area will be rectangular images, which can only be partitioned into 5 polygons, corresponding to the 5 points of the password of *PassPoint*. For this reason, at least 4 of these Voronoi polygons can remain open, but if the open polygons are neglected to avoid the edge effect, only one polygon at most should remain, and if the number of sides of the closest Voronoi polygons were estimated, the center would be the risk of being an open polygon. Therefore, the edges of the images will be taken into account when estimating the number of sides to obtain closed Voronoi polygons.

4.2 Design of the Experiment

The analyzes are performed for two image sizes, 800×480 and 1366×768 pixels, as they are the most common on mobile phones and computers, respectively. For each image, two databases of 300 random graphic passwords (R.G.P.) each were generated in *PassPoint*, with intensity $\lambda = 1.3021 \times 10^{-5}$ and $\lambda = 4.7660 \times 10^{-6}$, respectively, for a total of 1,500 closed Voronoi-Poisson polygons in each case. The estimation of the distribution of the number of sides of the Voronoi polygons was done without taking into account the order in which the pixels are selected by the user. The experiments carried out were developed in the R2018a version of Matlab. For each password in each database, the image was divided into the 5 Voronoi polygons, corresponding to the password points, and the number of sides, N , of each of the 1,500 polygons obtained in the 300 passwords of the 5 points each. The results are shown below.

4.3 Expected Value of the Number of Sides of the Voronoi Polygons in the Graphical Passwords of *PassPoint*

By partitioning the selected images into Voronoi polygons, with their corresponding random database, and counting the number of sides N of the Voronoi polygons associated with each of the pixels in the password, Table 2 is presented. The first database (DB.1) belongs to the image with size 800×480 pixels and to the second database (DB.2) the most frequent dimension in computers. These table is organized by the numbers of sides from highest to lowest absolute frequency to show

the numbers of sides that are more (4 and 5), less (6 and 3) and not (7) significant, which can be observed in Fig. 2. Using said table and figure, it is also possible to perceive the fit between the estimated distributions, due to the overlap between both databases provided by the accumulated frequency and the Fig. 2.

Table 2. Observed frequencies of the occurrence of the number of sides in the DB. 1 and 2 respectively.

| N | Frequencies DB.1 | Frequencies DB.2 | Relative F.(R.F.) DB.2 | 1-Cum.R.F DB.2 |
|-------|------------------|------------------|------------------------|----------------|
| 4 | 648 | 660 | 0.44000 | 0.56000 |
| 5 | 595 | 605 | 0.40333 | 0.15667 |
| 6 | 148 | 131 | 0.08733 | 0.06934 |
| 3 | 100 | 96 | 0.06400 | 0.00534 |
| 7 | 9 | 8 | 0.00534 | 0.00000 |
| Total | 1500 | 1500 | 1 | – |

The expected values and the variances associated with the random variable (number of sides) for DB.1 and DB.2 are, $E[N] = 4.5453$ and $E[N] = 4.53$, and $V[N] = 0.6147$ and $V[N] = 0.5838$ respectively. Note that both expected values are approximately 4.5 and the variances approximately 0.6 despite the differences between the intensities.

4.4 Evaluation of the Effectiveness of the Number of Sides of the Voronoi Polygons to Detect Clustering in the Graphical Passwords of *PassPoint*

In this subsection, a test based on the number of sides of the Voronoi polygons is proposed to detect grouping of points in the graphical passwords of *PassPoint*.

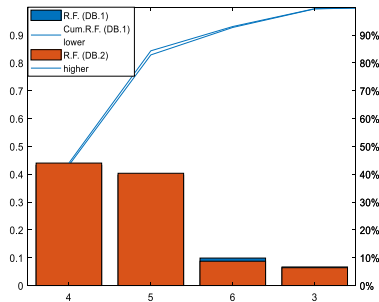


Fig. 2. Pareto diagram of the most and least significant numbers of sides observed in DB.1 and DB.2 respectively.

4.4.1 Proposal of a Randomness Test Based on the Number of Sides of the Voronoi Polygons in *PassPoint*

The following hypotheses are proposed: H_0 : The password points have been selected at random and H_1 : Otherwise, with test statistic given by the number of sides of the Voronoi polygons generated by the password selected by the user. The critical region defined from the numbers of sides that appear less frequently in graphical passwords whose points are random. There are 3 possible options: CR.1: $\{N > 6\}$, CR.2: $\{N = 3\} \cup \{N > 6\}$, CR.3: $\{N = 3\} \cup \{N > 5\}$.

4.4.2 Evaluation of the Effectiveness of the Test

To evaluate the effectiveness of the proposed test, type I and type II errors were measured.

Table 3. Probability estimated in DB.2.1, that the side number belongs to the critical region (*CR.*) under the hypothesis of randomness.

| Acceptance region | H_0 rejection region (<i>CR.</i>) | $\hat{p}(N \in CR. H_0)$ DB.2.1 |
|-------------------|---------------------------------------|--------------------------------------|
| $3 \leq N \leq 6$ | $\{N > 6\}$ | 0.00667 |
| $4 \leq N \leq 6$ | $\{N = 3\} \cup \{N > 6\}$ | 0.07934 |
| $4 \leq N \leq 5$ | $\{N = 3\} \cup \{N > 5\}$ | 0.18267 |

In Table 3, due to the adjustment of the distribution of the number of sides between DB.1 and DB.2, only a new database (DB.2.1) of 300 random graphical passwords was generated in an image with a size of 1366×768 pixels to estimate the type I error.

Note that since each graphical password is made up of 5 Voronoi polygons, it may be the case that a graphical password contains 0, 1, or more than 2 or more polygons whose number of sides belongs to the reject region.

As for the decision criteria, the graphical password selected by the user does not follow a random pattern if in the Voronoi polygons generated by it, there is at least a polygon with the number of sides that belongs to the rejection region. The graphical password follows a random pattern if all the side numbers of the Voronoi polygons generated do not belong to the rejection region, or they all belong to the acceptance region.

A new database (DB.3) of 300 grouped graphical passwords (G.G.P.) was generated in one sixteenth of the image of size 1366×768 pixels, and the proposed test was applied to each of the passwords, results shown in Table 4.

As can be seen in Table 4, for each of the 3 rejection regions, the proportion of passwords with grouped points that are rejected by the proposed test is very small. The highest effectiveness is obtained for CR.3, with a 53% rejection, which is still insufficient.

Table 4. Number of rejected graphical passwords (G.G.P.) observed in DB.3.

| Rejection region of H_0 | Number and proportion of G.G.P. detected in DB. 3 |
|----------------------------|---|
| $\{N > 6\}$ | 8/300 = 0.0266 |
| $\{N = 3\} \cup \{N > 6\}$ | 85/300 = 0.2833 |
| $\{N = 3\} \cup \{N > 5\}$ | 159/300 = 0.5300 |

Comparison of the Histograms of the Number of Sides for DB.2 and DB.3

The low effectiveness of the proposed test is also explained by the overlap of both distributions (by means of the green) illustrated in the following graph (Fig. 3).

4.5 Distribution and Evaluation of the Entropy of the Number of Sides of the Voronoi Polygons in the Graphical Passwords of *PassPoint*

In this subsection the fit between the estimated probabilities of the number of sides of the Voronoi polygons is measured for the random databases (Table 2) and the 54 theoretical distributions which supports the EasyFit 5.6 program [42, 43], with some of them for various parameter sets for a total of 61 distributions. This program allows you to automatically fit the distributions to the sample data and select the best model in a few seconds.

For the distribution of the number of sides of the polygons of the graphical passwords formed by 5 random points and contained in DB.1 and DB.2, sets of parameters other than those of Tanemura are obtained in [46] for a generalized gamma function of 3P, $\hat{\alpha}_1 = 55.219$, $\hat{\beta}_1 = 0.10566$, $\hat{\gamma}_1 = -1.2878$ and $\hat{\alpha}_2 = 59.359$, $\hat{\beta}_2 = 0.09934$, $\hat{\gamma}_2 = -1.3664$, respectively. When the Kolmogorov-Smirnov, Anderson-Darling and χ^2 tests are performed, they are rejected with

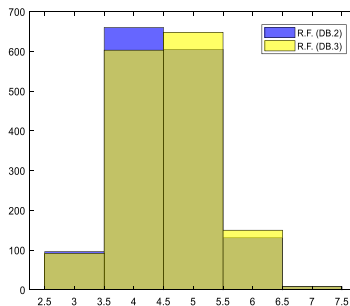


Fig. 3. Histograms of the observed frequencies of the number of sides of the Voronoi polygons in the random (DB.2) and clustered (DB.3) graphical passwords. (Color figure online)

significance levels $\alpha = \{0.2, 0.1, 0.05, 0.02, 0.01\}$ and a p -value = 0. The test statistics obtained for DB.1 were 0.24819, 103.22 and 720.95, respectively, and those of DB.2, 0.25289, 108.88 and 710.42, respectively. This result differs from those obtained in [23] and [46]. Even these distributions did not fit any of the 60 remaining distributions that this software brings by default.

4.5.1 Estimation of Entropy

In graphical authentication, entropy is used to measure the security of the password. In [4,29] they use the formula for the calculation: $H(x) = N \log_2(|L| |O| |C|)$, where N is the number of runs, L the Locus alphabet, O the target alphabet and C the alphabet color. For this they assume equiprobable passwords and maximum entropy and illustrate in [4] a comparison between some graphical authentication systems, including the *Passpoint*. But, dictionary attacks on these systems do not traverse the equally likely (randomly distributed) passwords, as they restrict their search space by selecting the most probable passwords. Therefore, their approach does not measure resistance against dictionary attacks with non-equiprobable password traversing.

To measure the level of uncertainty of this characteristic in random and grouped graphical passwords in *PassPoint*, the entropy will depend on the probabilities of the estimated number of sides. The entropy of said characteristic was estimated using parametric, non-parametric and semi-parametric estimators. A description and comparison of these estimators can be seen in [12]. In the background, when they calculate the entropy associated with the number of sides they only use the maximum likelihood estimator (ML) [6,7,31], although it is known to be a biased estimator. Also in [6,7] they call it Voronoi entropy, but this name is not correct because the entropy they used to measure the information is known as Shannon entropy. In this work, other estimators are used for a better estimation, the parametrics calculated were the Bayesian estimators Jeffreys, Laplace, Schürman-Grassberger, Minimax and finally the semi-parametric Shrink estimator [1].

To find the entropy estimators in a sample of 300 graphical, random and grouped passwords, these estimators were calculated for each of the passwords in the size image, 1366×768 pixels. The 5 different probability distributions associated with the side numbers that appear in the 300 passwords for the two databases coincide, and therefore their entropies. Although only the probability distributions shown in the Table 5 appeared, this does not mean that the following probabilities [0; 1; 0; 0; 0] and [1/5; 1/5; 1/5; 1/5; 1/5] might not be possible.

Table 5. Frequencies of appearance (FA) of the distributions in each type of database (random (DB.2) and clustered (DB.3)) and the entropy estimators (calculated in bits) associated with the number of sides of the 300 graphical passwords in each DB.

| \hat{P}_N | $F.(R.G.P)$ | $F.(G.G.P)$ | \hat{H}^{ML} | \hat{H}_{JEF}^{Bayes} | \hat{H}_{LAP}^{Bayes} | \hat{H}_{SG}^{Bayes} | $\hat{H}_{Minimax}^{Bayes}$ | \hat{H}_{Shrink}^{Bayes} |
|---|-------------|-------------|----------------|-------------------------|-------------------------|------------------------|-----------------------------|----------------------------|
| [1/5; 4/5; 0; 0; 0] | 0.1033 | 0.0767 | 0.7219 | 1.6879 | 1.9610 | 1.3153 | 1.6407 | 1.1409 |
| [2/5; 3/5; 0; 0; 0] | 0.4533 | 0.4300 | 0.9710 | 1.8228 | 2.0464 | 1.5051 | 1.7832 | 1.3521 |
| [2/5; 2/5; 1/5; 0; 0] | 0.1100 | 0.0933 | 1.5219 | 2.0419 | 2.1710 | 1.8530 | 2.0187 | 1.7600 |
| [1/5; 1/5; 3/5; 0; 0] | 0.1500 | 0.2333 | 0.9503 | 1.9628 | 2.1219 | 1.7396 | 1.9348 | 1.6331 |
| [1/5; 1/5; 1/5; 2/5; 0] | 0.1833 | 0.1667 | 1.3710 | 2.1819 | 2.2906 | 2.0875 | 2.1703 | 2.0409 |
| Total | 1.0000 | 1.0000 | - | - | - | - | - | - |
| \hat{H}_{max} | 2.2906 | 2.2906 | 1.5219 | 2.1819 | 2.2906 | 2.0875 | 2.1703 | 2.0409 |
| \hat{H}_{min} | 0.7219 | 0.7219 | 0.7219 | 1.6879 | 1.9610 | 1.3153 | 1.6407 | 1.1409 |
| $\hat{H}_{max} - \hat{H}_{min}$ | 1.5687 | 1.5687 | 0.8000 | 0.4940 | 0.3296 | 0.7722 | 0.5296 | 0.9000 |
| $\frac{\hat{H}_{max} - \hat{H}_{min}}{2}$ | 1.5063 | 1.5063 | 1.1219 | 1.9349 | 2.1258 | 1.7014 | 1.9055 | 1.5909 |

Table 5 allows us to visualize that in general, the maximum estimated value is obtained for the Laplace estimator, $\hat{H}_{LAP}^{Bayes} = 2.2906$, this value being close to the maximum value of the entropy $H_{max} = 2.3219$, for $k = 5$ categories. Also in this table there are values corresponding to the ML estimator that are “close” relatively to the values 1.65 and 1.71 respectively, but these values were calculated for points randomly distributed in [7] and [31] using the estimator ML, $H(P_N) = -\sum P_N \ln P_N$.

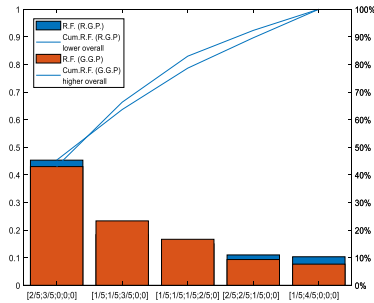


Fig. 4. Pareto diagram of the probability distributions of the number of sides of the Voronoi polygons that appear in the graphical random (DB.2) and clustered (DB.3) passwords.

The probabilities with which these estimators appear only differ significantly in the number of sides with probability [1/5; 1/5; 3/5; 0; 0], since this configuration is more frequent in grouped passwords. Therefore, due to the small sample and because they correspond to approximate probabilities to the entropy estimators, associated with the number of sides of the random graphical passwords and grouped with equal probability, it is not possible to distinguish whether the points are grouped or randomly distributed in said image, as shown in Fig. 4 by

the overlap between the probability distributions of the number of sides, which coincides with the result of [6], where they used the maximum plausible estimator to detect symmetry and not it worked. However, in [6] they later used the ML estimator associated with the length of the sides and it was able to distinguish symmetry.

5 Analysis of Results

An important result in this section is that the expected value continues to be approximately 5, regardless of the fact that the points are grouped in a certain region of the study area. Therefore, it is concluded that the characteristic “number of sides of Voronoi polygons” is not able to detect clustering in the graphical passwords of *PassPoint*, on criteria based on expected value and entropy, due to the similarity between the distributions of the number of sides in graphical passwords. Randomized and grouped in *PassPoint*, and due to having a small sample size, since only 5 observations are required (5 polygons in each password).

In our scenario, 300 sets of 5 Voronoi Poisson polygons were simulated in two images of sizes 800×480 and 1366×768 pixels, for an intensity $\lambda = 1.3021 \times 10^{-5}$ and $\lambda = 4.7660 \times 10^{-6}$ respectively. Previous works were generated in a set of the plane 100, 200 and in several iterations of 100,000 to 1,000,000 Voronoi Poisson polygons, with densities $\lambda = 0.16$, unknown and from $\lambda = 178$ until $\lambda = 1,782$ respectively. The known antecedent with the highest number of polygons generated was 208,969,210, but its intensity is unknown. Unlike in previous works, the expected value of the number of sides of the Voronoi polygons in *PassPoint* is 5. In *PassPoint* the number of estimated sides varies in a range from 3 to 7, while its range in previous works is from 3 to 13 by simulation and from 3 to 9 according to the explicit expression of Calka in 2002 [10]. The estimated distribution associated with the number of sides (or vertices) of the Voronoi polygons in previous works was approximated to a generalized three-parameter Gamma distribution, while its distribution in *PassPoint* has not been able to fit any known distribution. The value of the ML estimator associated with the probability distribution of the number of sides of Voronoi Poisson polygons for 200 points in a set of the plane is relatively close to some of the values in *PassPoint*; however, the calculation of the estimators differ on the basis of the logarithms.

6 Conclusions

In this work, the behavior of the number of sides of the Voronoi polygons generated by the graphical passwords of the *PassPoint* graphical authentication system was investigated. Its distribution was estimated, which could not be adjusted to any of the known distributions that the EasyFit program brings by default, including the generalized three-parameter gamma. It was obtained that the expected value of the number of sides of the Voronoi polygons was 5, regardless of the sizes studied. Therefore, the expected value of the number of

sides of the Voronoi polygons depends on the number of polygons generated in a study region, and therefore on the intensity, said result differs from that of the antecedents. In the studied scenario, the number of estimated sides varies between 3 and 7, not coinciding with the simulations of the antecedents, in which it varies from 3 to 13. Based on this distribution, a test was proposed, based on the expected value of the number of sides of the polygons, to detect weak graphical passwords formed by grouped points. The effectiveness of the proposed test was evaluated and it was concluded that it is not efficient for the detection of weak graphical passwords of *PassPoint* formed by 5 grouped points. The entropy of the distributions of the number of sides of the Voronoi polygons in random graphical passwords and in weak graphical passwords formed by grouped points was estimated. No significant differences were detected in the value of both entropies. It is concluded that the characteristic number of sides is not effective for the detection of weak graphical passwords of *PassPoint* formed by 5 grouped points. Future work will evaluate the ability of other features of Voronoi polygons, such as the perimeter of a Delaunay triangle and the length of one side of a Voronoi polygon (using the Voronoi entropy associated with this feature) to detect a clustering pattern in *PassPoint*.

References

1. Altay, G., Kurt, Z., Aydinl, N.: Comprehensive review of association estimators for the inference of gene networks. *Turkish J. Electr. Eng. Comput. Sci.* **24**(3), 695–718 (2016)
2. Aurenhammer, F., Klein, R.: Voronoi diagrams. In: *Handbook of Computational Geometry*, pp. 201–290. Elsevier Science, Amsterdam (2000)
3. Baddeley, A., Rubak, E., Turner, R.: *Spatial Point Patterns: Methodology and Applications with R*. CRC Press, Boca Raton (2015). ISBN-13:978-1-4822-2021-7
4. Bhanushali, A., Mangue, B., Vyas, H., Bhanushali, H., Bhogle, P.: Comparison of graphical password authentication techniques. *Int. J. Comput. Appl.* **116**(1), 0975–8887 (2015)
5. Boots, B.N.: *Voronoi (Thiessen) Polygons*. Published by Geo Books, ISSN 0306 6142 (1996)
6. Bormashenko, E.: Characterization of Self-Assembled 2D Patterns with Voronoi Entropy (2018). <https://doi.org/10.3390/e20120956>
7. Bormashenko, E., Legchenkova, I., Frenkel, M.: Symmetry and Shannon Measure of Ordering: Paradoxes of Voronoi Tessellation (2019)
8. Brakke, K.A.: 200,000,000 Random Voronoi Polygons (2015)
9. Brehcist, J.L., Herrera, J.: Mejoras de un sistema de contraseñas gráficas, *Universitat Autònoma de Barcelona, Máster interuniversitario de Seguridad de las tecnologías de la información y las comunicaciones* (2014)
10. Calka, P.: The explicit expression of the distribution of the number of sides of the typical Poisson Voronoi cell. Preprint of LaPCS, 02 Feb 2002
11. Chiu, S.N.: Spatial point pattern analysis by using Voronoi diagrams and Delaunay tessellations - a comparative study. *Biometr. J.* **45**(3), 367–376 (2003)
12. Contreras, L., Legón, C.M., Madarro, E., Socorro, R.: Estimación de la entropía en sucesiones aleatorias cortas de bytes y bits, Tesis presentada en opción del título de Máster en Ciencias en la Facultad de Matemática y Computación, Instituto de Criptografía, Facultad de Matemática y Computación, Universidad de La Habana (2020)

13. Crain, I.K.: Monte-Carlo simulation of the random Voronoi polygons-preliminary results. *Search* **3**(5), 220 (1972)
14. Crain, I.K.: The Monte Carlo generation of random polygons. *Comput. Geosci.* **4**, 131–141 (1978)
15. Dobrin, A.: A review of properties and variations of Voronoi Diagrams. Whitman College (2005). 10.1.1.453.9156
16. Edwards, R., Mardia, K.V., Puri, M.L.: Analysis of central place theory. *Bull. Int. Stat. Inst.* **47**, 93–110 (1977)
17. Ferraro, M., Zaninetti, L.: On the statistics of area size in two-dimensional thick Voronoi diagrams. *Physica A Stat. Mech. Appl.* **391**(20), 4575–4582 (2012)
18. Gelfand, A.E., Diggle, P.J., Fuentes, M., Guttorp, P.: *Handbook of Spatial Statistics*, CRC Press, Boca Raton (2010). ISBN 978-1-4200-7287-7
19. Goodman: *Handbook of Discrete and Computational Geometry*, 3rd edn. CRC Press (2017). LCCN 2017017843, ISBN 9781498711395
20. Hayen, A., Quine, M.: The proportion of triangles in a Poisson-Voronoi tessellation of the plane. *Adv. Appl. Prob. (SGSA)* **32**, 67–74 (2002a)
21. Hayen, A., Quine, M.: Calculating the proportion of triangles in a Poisson-Voronoi tessellation of the plane. *J. Stat. Comput. Simul.* **67**, 351–358 (2002a)
22. Herrera, J.A., Legón, C.M., Piñero, L.R., Sosa, G., Rojas, O.: Effectiveness of spatial randomness test in detection of weak graphical passwords in passpoint. In: 4th EAI Internacional conference on Computer Science and Engineering in Health Services (COMPSE 2020) (2020)
23. Hinde, A.L., Miles, R.E.: Monte Carlo estimates of the distributions of the random polygons of the Voronoi tessellation with respect to a Poisson process. *J. Stat. Comput. Simul.* **10**, 205–223 (1980)
24. Icke, V., Van de Weygaert, R.: Fragmenting the universe. *Astron. Astrophys.* **184**(1–2) (1987). ISSN 0004–6361
25. Kenkel, N.C., Hoskins, J.A., Hoskins, W.D.: Edge effects in the use of area polygons to study competition. *Ecology* **70**(1), 272–274 (1989)
26. Kirovski, D., Jodic, N., Roberts, P.: Click Passwords, Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA (2020)
27. Komanduri, S., Hutchings, D.R.: Order and entropy in picture passwords. In: *Graphics Interface Conference 2008*, Windsor, Ontario, Canada (2008)
28. Kumar, S., Kurtz, S.K.: Properties of a two-dimensional Poisson-Voronoi tessellation: a Monte-Carlo study. *Mater. Charact.* **31**(1), 55–68 (1993)
29. Lashkari A.H., Salleh, R.: A new algorithm for graphical user authentication based on rotation and resizing, A thesis submitted for the master of Computer Science in Data Communication and Computer Networking, Faculty of Computer Science and Information Technology, University Malaya (2010)
30. Lembach, S., Gebert, J.R.: *Voronoi and Delaunay diagrams*, Technische Universitat Munchen (2010)
31. Limalle, A.V., Narhe, R.D., Dhote, A.M., Ogale, S.B.: Evidence for convective effects in breath figure formation on volatile fluid surfaces. *Phys. Rev. Lett.* **76**(20), 3762–3765 (1996)
32. Liu, B., Meng, Q., Holstein, H.: Point pattern matching and applications - a review. *IEEE Xplore* (2003). <https://doi.org/10.1109/ICSMC.2003.1243901>
33. Miles, R.E.: On the elimination of edge-effects in planar sampling. In: Harding, E.F., Kendall, D.G. (eds.) *Stochastic Geometry*, pp. 228–247. Wiley, London (1970)
34. Møller, J.: *Lectures on Random Voronoi Tessellations*. LNS. Springer, New York (1994). <https://doi.org/10.1007/978-1-4612-2652-9>

35. Muche, L.: Distributional properties of the three-dimensional Poisson Delaunay cell. *J. Stat. Phys.* **84**, 147–167 (1996)
36. Mumm, M.: Voronoi diagrams. *Math. Enthusiast* **1**(2) (2004). Article 4
37. Nakoinz, O., Knitter, D.: *Modelling Human Behaviour in Landscapes*. Springer, Switzerland (2016). <https://doi.org/10.1007/978-3-319-29538-1>
38. Okabe, A., Boots, B., Sugihara, K., Chiu, S.N.: *Spatial tessellations: Concepts and Applications of Voronoi Diagrams* (2000). ISBN 0-471-98635-6. British Library Cataloguing in Publication Data
39. Ozcan, M., Yaman, U.: A continuous path planning approach on Voronoi diagrams for robotics and manufacturing applications. In: *29th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM 2019)*, 24–28 June, Limerick, Ireland (2019)
40. Rittenhouse, R.G., Chaudry, J.A., Lee, M.: Security in graphical authentication. *Int. J. Secur. Appl.* **7**(3), 347–356 (2013)
41. Rodríguez, O., Legón, C.M., Socorro, R.: Seguridad y usabilidad de los esquemas y técnicas de autenticación gráfica, revista cubana de Ciencias Informáticas, vol. 12, no. Especial UCIENCIA, 13–27 (2018)
42. Schittkowski, K. (2000): *EASY-FIT: A Software System for Data Fitting in Dynamical Systems*
43. Schittkowski, K.: *Data Fitting in Dynamical Systems with EASY-FIT -User's Guide* (2002)
44. Schmid, C., Leitner, M.: Monte-Carlo simulation of two-dimensional grain growth (2011)
45. Snibbe, S.S., Tamassia, R.: Introduction to Voronoi diagrams. In: *Computational Geometry, C.S. vl. 252* (1993)
46. Tanemura, M.: Statistical distributions of Poisson Voronoi cells in two and three dimensions. *FORMA-TOKYO* **18**(4), 221–247 (2003)
47. Tico, M., Rusu, C.: *Point Pattern Matching using a Genetic Algorithm and Voronoi Tessellation*. Tampere University of Technology, Signal Processing Laboratory (1998)
48. Wiedenbeck, S., Waters, J., Birget, J.C., Brodskly, A., Memon, N.: Passpoints: design and longitudinal evaluation of a graphical password system. *Int. J. Hum Comput Stud.* **63**(1–2), 102–127 (2005)