Evaluating Mobile Phone Handoff Behavior using Chi-square Statistical Test

Saeed A. Rajput,
Senior Member IEEE
Farquhar College of Arts & Sciences, Nova Southeastern University, Fort Lauderdale, FL, 33134
rajput@nova.edu

A. S. Pandya and S. Saxena
Dept. of Computer Science and Engineering
Florida Atlantic University, Boca Raton, FL 33431
{pandya,ssaxena}@fau.edu,

Steve Ostroff
Motorola, Inc., Plantation, FL 33322
Steve.Ostroff@motorola.com

Abstract

Mobile networks reuse frequency bands based on a color map to increase the capacity of the network. A handoff should occur when a mobile unit moves from the influence of one base station with weaker signal into another's that has stronger signal. Handoff behavior of all units is an important factor in quality of service of a mobile phone service. Handoff decisions, also called mobility decisions, are made by mobile phone based on the observed power from base stations. Premature, delayed or exceedingly sensitive decisions are considered poor decisions. Excessive poor decisions result in degradation of service quality in otherwise a healthy mobile system. Conventional research focuses on improving hand-off algorithms. Most of the published work on verification of effectiveness of handoff algorithms is analytical or focuses on data collected under pristine laboratory conditions. A unit that makes good mobility decisions, theoretically or in the laboratory, may not behave as expected in the real world however. We propose a process of evaluating hand off behavior using large amount of diagnostic phone data collected in the real world that is used for identification of adverse trends or aberrant behavior of various models. In this paper, we discuss a chi-square statistical test to evaluate the performance of specific mobile unit model by comparing the behavior of a test mobile unit against a well-established behavior profile. If the behavior of the test model deviates significantly from the well-established profile, it is considered deficient in its handoff behavior that deserves further analysis. The test was developed in such a way that a large amount of units can quickly be tested. The same test can be used to compare performance of all mobile phones in one region to performance of same mobile phones in other regions. Furthermore, our test is useful in determining difference of handoff behavior when the mobile units are moving in opposite directions.

1. Introduction

Since the invention of wireless communications [i], innovations have continuously pushed the limit of capacity of wireless systems. First, Frequency Division Multiplexing (FDM) was used to transmit and receive multiple signals simultaneously. Then Time Division Multiplexing (TDM) was used to further expand the capacity of communications. [ii]. Ever increasing need for wireless call traffic on limited electromagnetic spectrum or frequency band led to the development of the cellular radio networks [iii]. It increases the number of simultaneous conversations possible (called capacity) in a given mobile radio telephone system by frequency reuse [iii]. In cellular networks, numerous lower-power transceivers, each with shorter coverage, are strategically deployed to cover a large geographic area. Frequency reuse factor expresses the rate at which a given frequency can be used in a given network. Several low power transceivers are part of a base station that services a cell. The base station also contains a control unit that manages continuity of individual phone calls. Multiple frequency bands are assigned to each cell, so the base station is capable of communicating with mobile phones in that cell on any of those frequencies. In a given cellular architecture, the number of frequency bands assigned to a cell depends on the traffic expected in that cell. Adjoining cells are assigned different group of frequencies to avoid interference and crosstalk. However, cells that are sufficiently far apart can reuse the same frequencies since radio signals strength diminishes with distance [iii]. To achieve this, frequencies are usually bundled into colors. Thus, the problem of frequency allocation maps to the map-coloring problem [iv,v]. In a real cellular network, at any given instance, a number of mobile units are moving in a cell, and they actively communicate with the base station. Several channel assignment strategies are discussed in the conventional research publications, for instance [vi] discusses an aggressive channel allocation scheme to reduce call loss due to failed hand over requests and call blocking in a multiple call hand-off context.

Each base station is connected to a mobile switching center, which serves multiple base stations. The mobile switching center routes the calls depending on the location of the mobile unit, assigns voice channel to each call, performs handoffs, and monitors the call for billing information [iii]. We will call this mechanism mobile call management.

978-1-4244-1884-8/08/$25.00 ©2008 IEEE

When a mobile unit moves from coverage area of one base station to another it is said to have *crossed the cell boundaries*. The mobile unit must switch the traffic channel assigned to the old base station to the new base station as it crosses over its cell boundary. This process is called handoff and is performed ubiquitously. The Received Signal Strength Indicator (RSSI) gets weaker as the mobile unit moves away from the base station. When a neighboring site is stronger than the serving/current cell, the mobile unit requests a handover to another site [iii]. After the handover, the RSSI typically sees significant improvement. Cell dragging may occur when a mobile handset moves a considerable distance into the neighboring cell area without making a handoff, resulting in an increased level of system interference [vii]. A handoff scheme that utilizes two adaptive algorithms in combination; one using a relative threshold and the other an absolute threshold, has been proposed in [vii]. This scheme aims at minimizing cell dragging. Other works [viii and ix] have proposed models to configure cellular networks based on subscriber mobility between cells. In [x] Magnusson and Olofsson have proposed an algorithm to dynamically determine neighboring cell lists i.e. handover candidates and their associated broadcast control channels for each cell in the system. Bansal et.al [xi] propose using knowledge of the cell terrain, the mobile trajectory, and the vehicular movements in a cellular network to predict handoff rates.

In this paper, we do not try to evaluate or compare the performance of various handoff algorithms, but evaluate the performance of various handsets models. Little has been published to-date that provides methods of analyzing the real world performance handsets from the user's perspective. In this paper, we develop a comprehensive statistical test that measures the degree to which any handset is implementing the handoff design and it evaluates the handset's performance independent of the handoff algorithm implemented. The focus of our evaluation is on the behavior of the handset irrespective of what handoff algorithm the service is using.

In section 2, we provide the profiles of desire able and undesirable mobility behaviors. In section 3, we discuss our strategy in how the data is collected, and aggregated. We also provide our stochastic model and how the data is used to calculate the statistical chi-square test on our data. In section 4, we provide the results and show that we were able to successfully distinguish models that made poor mobility decisions with our approach. To protect the proprietary information, we have to use generic names for our models.

2. Mobility behavior

A normal decision consists of a handover to the new tower that provides the unit with higher RSSI (stronger signal) as a mobile phone moves from a weaker tower site to its neighborhood. This decision process is illustrated in Figure 1. Position *1* in Figure 1 depicts the behavior of a mobile unit while it is moving from one cell to another cell. At position *1,* it is within the coverage region of Base station A that has assigned color code ***X*.** It is moving towards Base station B that uses color code *Y*. At position 2 it moves away from A, it goes in the overlapping region between the two base stations. In this region the RSSI from two stations are close in values. To avoid excessive handoffs in this region the mobile phones are designed to retain connection with the previous tower until the signal from tower B is at least 6dB stronger.

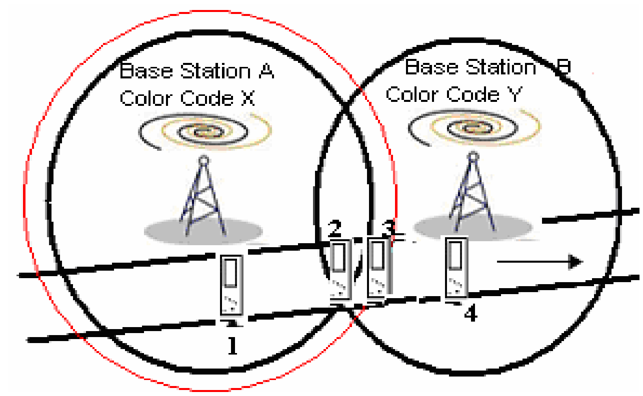

Figure 1 Mobility Behavior of handset

When the mobile moves to position *3* it is closer to base station B than the base station A. The handset at this point is expected to perform a handoff to cell site B with the new frequency. When the handset (position *4*) crosses the range of base station A, a normal phone switches the mobile communication with the base station B and performs a handoff such that it now uses color code Y of cell site B. The base station A hands over the communication channel it allocated to communicate with the mobile to base station B so that it can establish the same communication with the mobile handset.

Under these conditions if the mobile fails to handoff, it is said to make a poor mobility decision. The poor mobility decision in this case would be to continue to use base station A while base station B's RSSI is higher than 6dB. This is called cell dragging where a mobile handset moves a considerable distance into the neighboring cell area B without making a handoff. This abnormal behavior of cell dragging mostly leads to poor audio quality during the call. It can also cause the call to be dropped. Real world is much more asymmetrical than what is depicted in Figure 1. Signals from multiple towers might not overlap and the signal may not decay in a non-increasing fashion due to geographical and physical conditions. Poor mobility decisions are made due to variety of reasons that may or may not depend on the quality of the handset. An ill-designed phone exhibits poor mobility decisions more frequently than normal units do.

3. Model and Strategy

In analysis the handoff performance of mobile phones using their relative performance in real world environment. In this process, first a "baseline unit" is selected that has a specific model which is generally known to behave well in the field. The unit belonging to a specific mobile phone model to be tested (called test model) is then compared to the baseline unit.

3.1. Data Acquisition

To collect real data for this analysis, the baseline units and/or the test models are mounted on a test vehicle and are connected to the test equipment. The vehicle is driven over a geographic route that is in a closed loop (called drive-loop, drive test loop or simply drive test) and the diagnostic data is recorded in the test logs. Hence, this data represents a live phone call in a real world condition. Such drive tests produce approximately one record every second. Each record includes latitude, longitude, color code and RSSI data captured at that instance while the phone is in the call. Color codes are used to distinguish carrier signal from its co-channel interference. A color code set consists of different integral numbers. Each number is allocated to one of the frequency sets that are geographically nearby to each other [xii]. The comparison between test model and baseline unit is made by comparing the data in these test logs. To make these tests more meaningful, data is analyzed separately for clockwise and counter-clockwise direction in the drive test loop. Using similar techniques, we can also identify regions of coverage or each tower based on RSSI data and determine significant deviation of a mobile phone's RSSI from neighboring mobile phones.

3.2. Data Aggregation

Our analysis is based on the "color code" and Global Positioning System (GPS) location. Using the GPS data, the drive test route is divided into 50m x 50m squares and color code information is aggregated. In a square, number of records collected is ranked based on a measure that is based on statistical techniques. The choice of 50m x 50m squares seemed to provide most consistent results for our analysis. In the early phase, we tried smaller grid size of 30m x 30m and larger grid size of 100m x 100m. Smaller grids did not have enough data in unit areas to provide significant sample sizes, while in the bigger grids data resolution reduced. One observation while analyzing smaller grids was that we had more valid grid points to use for comparison. Due to the motion of the car and the one-second delay between samples associated with moving at 60mph it was not feasible to choose squares smaller than 30m.

The aggregated data is then used to identify poor mobility decisions that are made by the mobile phone and those mobile phones that exhibit an aberrant behavior.

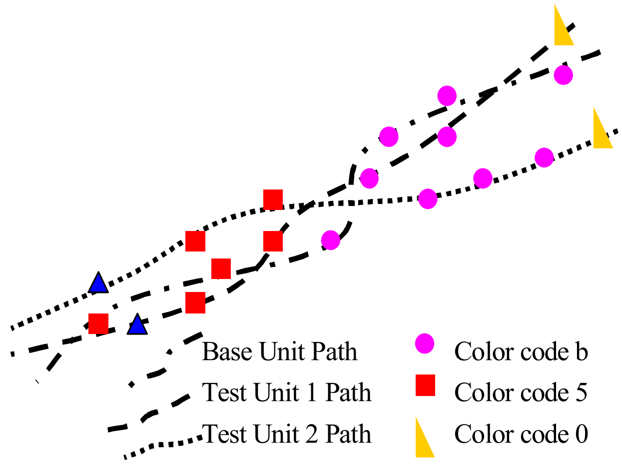

Figure 2 Color Codes observed for three drive tests

To aggregate the data for a cell we count the number of same color code samples for the same type of mobile phone in the square. The number of samples with the same color is then divided by the total number of observations in that square to obtain the frequency of that color. If we treat the color code as a random variable problem then the distribution of this random variable can be estimated by the color frequency distribution mentioned above. The underlying sample space for a given square for this random variable will have all the color codes.

Figure 2 shows a sample square in the geographic grid with three drive paths and the colors of the observations. The paths of the drive vehicle are indicated by dashed curves while the colored symbols indicate the color observations during those drives. We have used different symbols for each color.

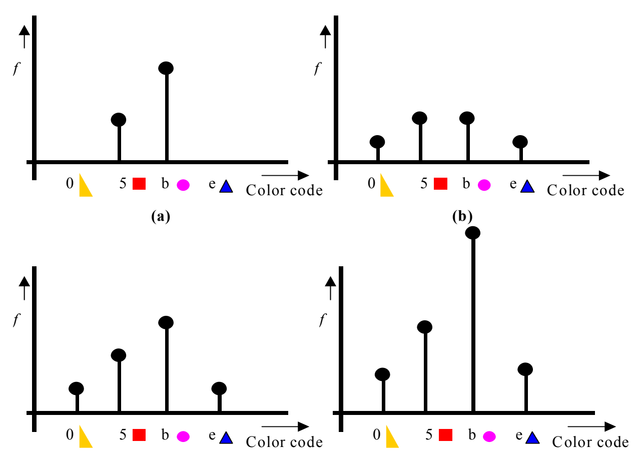

Figure 3 Frequency distribution for base unit (a), test model 1 (b), test model 2 (c) and aggregated frequency distribution for all three units (d).

To obtain the frequency distribution, we just count the number of times each color is observed for each drive, and divide it by the total number of observations on that drive. The resulting probability distributions are shown in Figure 3 (a) for base unit, Figure 3 (b) for test model 1, and Figure 3 (c) for test model 2. To obtain a more accurate distribution, we can also aggregate data from multiple test drives that belong to the same mobile phone model as illustrated by the distribution shown in Figure 3 (d).

3.3. Stochastic Model

The basic unit of our stochastic process is the square in the grid over which we are making observation. The basic experiment consists of two test drives through this square, one for the test model and the other for the base unit. We have already discussed the color code distribution. That corresponds to a color code random variable for a given mobile phone we will denote this random variable with $\boldsymbol{c_{i,j,t}}$, where *(i,j)* correspond to the coordinates of the square in the grid and *t* corresponds to the test ID of the particular drive. We can also treat phone unit as a random variable as well and represent it by $p_{i,j,t}$ In our model, this random variable has two values base unit and test model. The two random variables $\boldsymbol{c_{i,j,t}}$, and $p_{i,j,t}$ form a joint probability space for a given cell. Because there are 16 color codes in one cellular model and two units, we have 32 possible outcomes. For instance (base, red) is an outcome that indicates that the base unit is connected through the red frequency. The probability of this outcome is given by the relative frequency distribution observed for that color code and the mobile phone. Thus, this joint probability space can be organized as a contingency table. Table 1 is an example of a 2 x 16 contingency table for the two phones (baseline and test) and 16 color codes. The column total and row total are marginal totals.

Table 1 A 2 x 16 Contingency table for Color codes and phones

	Color Code																
	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f	*Total*
Model	0	0	30	38	0	0	0	0	0	0	0	0	0	0	0	0	*68*
Test	0	0	16	65	0	0	0	0	0	0	0	0	0	0	0	0	*81*
Total	*0*	*0*	*46*	*103*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	***149***

In our test, we want to determine if the selection of mobile phone affects the color code distribution. If the two mobile phones behave in a similar fashion, the color code distribution should be unaffected. However, if the mobile phone behaves differently then they have different color code distribution for that given square. That would imply that the color random variable is dependent on mobile phone random variable.

Hence, we can setup this as statistical Independence hypothesis testing problem. The Null Hypothesis is that the random variable color is independent of mobile phone selection.

H_0: Color code distribution is independent of mobile phone selection (i.e. $c_{i,j,t} \coprod p_{i,j,t}$)

H_1: The variables are dependent

To establish this independence, our approach is to calculate the Chi-Square test statistic for each square and we will reject all the squares in which we fail to reject the null hypothesis using the p-value test. We will only keep the squares in which the null hypothesis is successfully rejected. These are the squares where changing the two mobile phones provided different color code distributions.

In the next section we will compare the hand-off behavior of mobile phones in a square using one test drive for the base profile and one test drive for the test model. We compare the hand-off behavior of mobile phones using the aggregated data from multiple test-drives for the base profile and the test model in selected squares of our grid.

4. Results

In this section, we will first compare the handoff behavior using one test drive for the base profile and one test drive for the test model. To improve the accuracy of our results we will then compared data aggregated from multiple drives so that the number of samples in a single square is higher. The latitude and longitude coordinates of the square that we study here are 176 and 415 respectively.

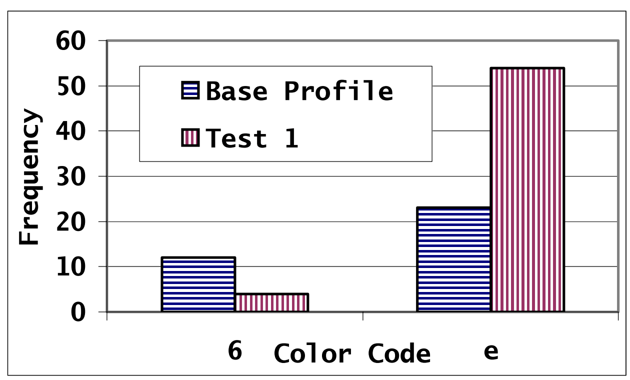

Figure 4 Test model 1 compared with the base profile using data from a single test drive, Chi-Square = 11.495, degree of freedom = 1 and p-value = 0.0007

Figure 4 shows data from when a single drive test data of base profile is compared with the test data of a single drive of out test model 1. In this data, there are two observed frequencies of colors *6* and *e* in this square so our contingency table will be of size 2 x 2). The Chi-square value for this comparison is 11.49544 and degree of freedom (DoF) is 1. At significance level 0.01 using the p-value 0.000698, we can infer that the handoff behavior for

Test 1 is significantly different from base profile and requires further inspection.

Similarly, we compare base profile with test model 2. Again we see the same two colors 6 and e but different frequencies as shown in Figure 5. At significance level 0.01 using the p-value 0.082346, we can infer that the handoff behavior for Test 1 is significantly similar to the base profile and should be a good mobile handset model.

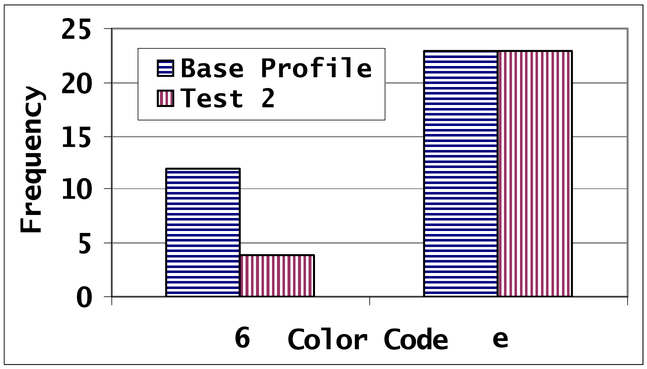

Figure 5 Test model 2 compared with the base profile using data from a single test drive, Chi-square = 3.018 degrees of freedom = 1 and p-value = 0.08235

For the same square we now compare multiple test-drives for base profile to multiple test-drives of test model 1, test model 2 and a handset moving in opposite direction. Our first comparison was base profile to test model 1(Figure 6). At significance level 0.01, we find the handoff behavior for Test 1 is significantly different from base profile and requires further inspection.

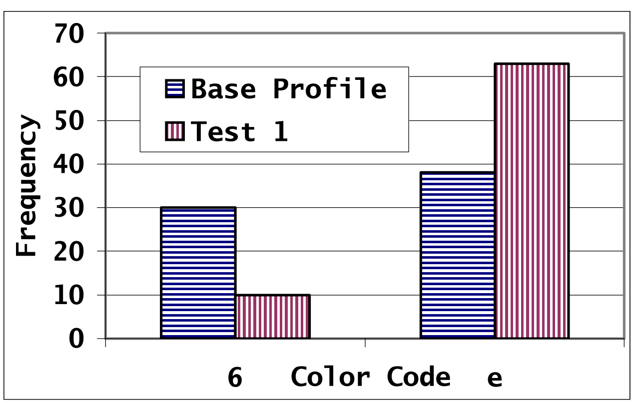

Figure 6 Test model 1 compared with the base profile using data aggregated from multiple test drives. Chi-square = 16.031, degree of freedom = 1 and p-value = 6.23x10-5

The comparison of multiple test-drives for the base profile and Test model 2 is described in Figure 7. Using the p-value 0.618374, we can clearly infer that test model 2 behaves similar to the base profile and this mobile handset model should be a good phone. We also conducted a comparison to analyze the affect of directionality on handovers in the same square.

Figure 8 shows the color code distribution for two handsets moving in opposite directions. At a significance level 0.01, we can infer that this test model moving in direction 0 has a handover behavior that is notably different from the base profile moving in direction 1.

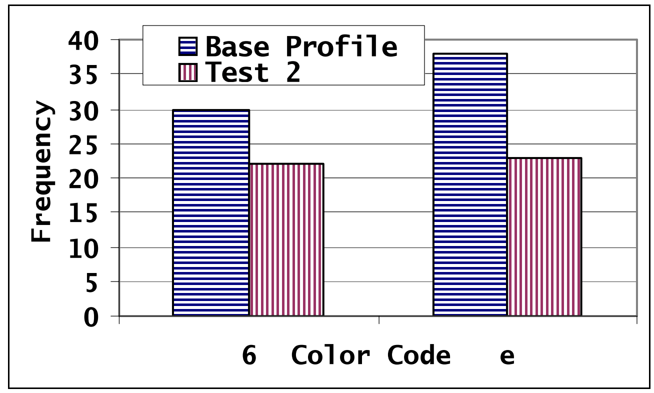

Figure 7 Test model 2 compared with the base profile using data aggregated from multiple test drives. Chi-square=0.248 Degree of freedom=1 and p-value = 0.6184

These results evidently show that Chi-Square Statistical Test is very powerful in comparing the handover behaviors for not only one test drive to another single test drive but also aggregated data of multiple test-drives. The comparison of a single test drive and aggregated data for the same test model yielded same results. The effects of directionality on handovers can also be examined using Chi-square Statistical Test.

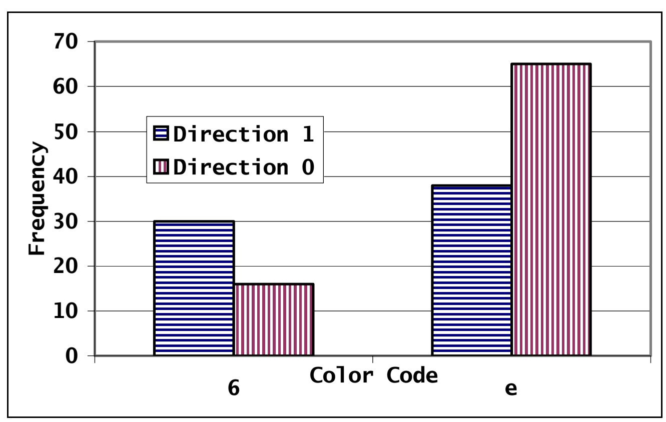

Figure 8 Comparison of same model driven in different directions using aggregated data from multiple drives. Chi-square=10.283, degree of freedom =1 and p-value = 0.001343

Our results are summarized in the Table 2. On this cell, we can clearly see that the behavior of test model 2 significantly different from the base profile. We also see that test model 2 is clearly acceptable.

Table 2 Summary of Chi-Square Value for Aggregated Drive Test Data

	Test Model 1	Test Model 2	Reverse Direction
Chi-Square	16.03097	0.248159	10.28259
DoF	1	1	1
p-value	6.231E-05	0.6183744	0.0013429

5. Conclusion

We have discussed that handoff behavior of all units in the real world is an important factor for high quality of service in a mobile phone system. In this paper, we have provided method to identify mobile phone models that frequently exhibit premature, delayed or exceedingly sensitive handoff decisions that are considered poor mobility decisions.

Unlike conventional research that either focuses on improving hand-off algorithms, or evaluates handoff algorithms analytically or uses data generated under controlled laboratory conditions, we have evaluated hand off behavior based on diagnostic data collected from mobile phones in the real world, while treating the mobile unit as a black box.

It is our experience that a unit that makes good mobility decisions may not behave as expected in the real world. Therefore, we have proposed treating the test mobile phone as a black box for our analysis to obtain dependable results. We have shown the feasibility of chi-square statistical test as a measuring mechanism for determining the hand-off behavior of mobile phones. We have shown that a contingency table can be developed based on observed frequencies of color in a small geographic region. This contingency table can then be used to calculate chi-square value using the approach of the independence hypothesis testing. The chi-square value can then be used to calculate the p-value that is used to accept or reject the independence hypothesis using a threshold. We have used this test to compare hand-off behavior in certain geographical regions. We have also shown that the measure can be used to establish difference of handoff behavior in different directions. Finally, we have shown that the analysis can be used to distinguish mobile phones based on their hand-off behavior.

6. Acknowledgements

The authors are grateful to all the people who made this study possible. We also want to thank Dr. Sam Hsu for the administrative support for the project, Max Haider for his ideas, Xiaoguang Sun for database scripts, Arjumand Siddiqui for initial investigations, Teerapat Therrawong, Pateep Kitbamrung, Rashid Jillani and Robert Stone for the support they provided in our Telecommunications and Computer Networking Lab, Florida Atlantic University and their contributions to this research.

7. References

[i] G. Marconi, "Wireless Telegraphic Communication: Nobel Lecture, December 11, 1909", *Nobel Lectures*, *Physics 1901-1921*, Elsevier Publishing Company, Amsterdam, 1967, pp. 196-222.pp 198, http://nobelprize.org/nobel_prizes/physics/laureates/1909/ marconi-lecture.pdf.

[ii] S.K. Shanmugan, *Digital and Analog Communication Systems*, Wiley, 2004, ISBN 9971511460

[iii] S. T. Rapport, *The Cellular Concept- System Design Fundamentals*, *Wireless Communications: Principles and Practice*, Ch 3, Prentice Hall, New Jersey 2002, pp. 57-58.

[iv] A. Cayley, "On the colourings of maps", *In Proc. of of Royal Geographical Society*, Vol. 1, 1879, pp. 259-261.

[v] K. Appel, W. Haken and J. Koch, "Every Planar map is Four Colorable", *Journal of Mathematics,* Illinois, December 1977, Vol. 21, pp. 439-567.

[vi] D. Zeghlache, "Aggressive Handover Algorithm for Mobile Networks", *In Proc. of 44th IEEE Conf. on Vehicular Technology*, Stockholm, June 1994, Vol. 1, pp. 87-90.

[vii] G. Senarath, A. Abu-Dayya and R. Matyas, "Adaptive Handoff Algorithms Using Absolute and Relative Thresholds for Cellular Mobile Communication Systems", *In Proc. of 48th IEEE Conf. on Vehicular Technology,* Ottawa, Ont., Canada, May 1998, Vol. 2, pp. 1603-1607.

[viii] B. Gavish and S. Sridhar, "The Impact of Mobility on Cellular Network Configuration", *Wireless Networks,* Springer, Netherlands, March 2001, Vol. 7, No. 2, pp. 173 - 185.

[ix] R. Sankar and N. Savkoor, "A Combined Prediction System for Handoffs in Overlaid Wireless Networks", *In Proc. of IEEE International Conf. on Communications,* Vancouver, BC, Canada, June 1999, Vol. 2, pp. 760 – 764.

[x] S. Magnusson, and H. Olofsson, "Dynamic neighbor cell list planning in a micro cellular network", *In Proc. of 6th IEEE International Conf. on Universal Personal Communications Publication,* San Diego, CA, , October 1997, Vol. 1, pp. 223-227.

[xi] D. Bansal, A. Chandra, R. Shorey, A. Kulshreshtha and M. Gupta, "Mobility models for cellular systems: cell topography and handoff probability", *In Proc. of 49th IEEE Conf. on Vehicular Technology,* Houston, TX, July 1999, Vol. 3, pp. 1794 - 1798.

[xii] Z. Lin, S. J. James, "Automatic color code (SAT) assignment method used in frequency planning for wireless networks", Ericsson Inc. ,Plano, TX, US Patent 6507742, November 11, 1999, http://www.patentstorm.us/patents/6507742-fulltext.html