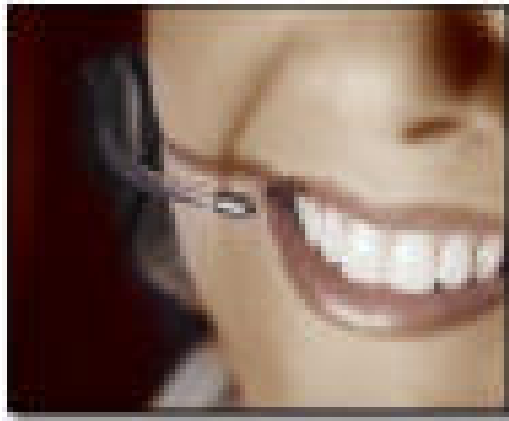




פרויקט שנתי בהנדסת תעשייה וניהול

ניתוב מבוסס מיומנות במוקדי

שירות טלפוני



מנחה:

פרופסור אבישי מנדלבאום

חברי הקבוצה:

אברהמי מירב

הכרי יחיאל

לוזון יוסף

רוזנברג ליהי

יולי 2004

תודות

במהלך פרויקט זה אספנו חומר רב בנושא ניתוב מבוסס מיומנות במוקדי שירות טלפוניים. נעזרנו במאמרים שונים וביקרנו במוקדי שירות שונים על מנת להבין לעומק את הנושא.

ברצוננו להודות למספר אנשים שעזרו לנו במהלך הפרויקט:

- מר רונן רומנו, מנהל צוות טלפוניה ומחשוב "פלאפון".
- מר תומר יהלום, שעזר לנו רבות בסימולצית מודל V.
- שרון פרנקו, גידי מקלף, אשר צוקר, איגור גולוד ואלה שרפשטיין מבנק הפועלים.

נרצה להודות עד מאוד למר איתי גורביץ' שעזר לנו רבות במהלך הפרויקט, וענה בשמחה על כל שאלה או בעיה בה נתקלנו.

מעל לכל, אנו רוצים להודות לפרופסור אבישי מנדלבאום אשר הנחה אותנו לאורך כל הפרויקט ופתח את עינינו לתחום הנדסת מערכות השירות בכלל, וניתוב מבוסס מיומנות בפרט.

תכולת הדיסק המצורף –

1. Presentation – המצגת המסכמת של הפרויקט כולו.
2. – V model
 - 2.1. results -
 - 2.1.1. results - total - vip=0.5reg.xls – סיכום התוצאות עבור המקרה בו קצב הגעת לקוחות ה-VIP שווה למחצית קצב הגעת הלקוחות הרגילים. כולל גליונות אלקטרוניים וגרפים.
 - 2.1.2. results - total - vip=2reg.xls – סיכום התוצאות עבור המקרה בו קצב הגעת לקוחות ה-VIP שווה לפעמיים קצב הגעת הלקוחות הרגילים. כולל גליונות אלקטרוניים וגרפים.
 - 2.1.3. results total vip=reg.xls – סיכום התוצאות עבור המקרה בו קצב הגעת לקוחות ה-VIP שווה לקצב הגעת הלקוחות הרגילים. כולל גליונות אלקטרוניים וגרפים.
 - 2.2. programs –
 - 2.2.1. main.cpp – תוכנית הסימולציה בשפת C++ המתאימה למקרה בו קצב הגעת לקוחות ה-VIP שווה לקצב הגעת הלקוחות הרגילים.
 - 2.2.2. v0.5reg.cpp – תוכנית הסימולציה בשפת C++ המתאימה למקרה בו קצב הגעת לקוחות ה-VIP שווה למחצית קצב הגעת הלקוחות הרגילים.
 - 2.2.3. v2reg.cpp – תוכנית הסימולציה בשפת C++ המתאימה למקרה בו קצב הגעת לקוחות ה-VIP שווה לפעמיים קצב הגעת הלקוחות הרגילים.
3. – N model
 - 3.1. results -
 - 3.1.1. data N-N.xls – קובץ הפלט לסימולציה כאשר ישנם N שרתים בכל קבוצת שרתים. מכיל את הפלט, הניתוח והגרפים.
 - 3.1.2. total.xls – מכיל את התוצאות והגרפים המסכמים של הקבצים המוזכרים ב-3.1.1.
 - 3.2. main.cpp – תוכנית הסימולציה בשפת C++ עבור מודל N.
 - 3.3. Annual project.doc – הפרויקט המלא.

תוכן העניינים

1.	מבוא.....	6
2.	מוקדי שירות טלפוניים – מהם?.....	8
2.1.	מוקדי שירות טלפוניים מהם?.....	9
2.2.	מידול מוקדי שירות.....	11
2.3.	מדדים לבחינת רמת השירות.....	12
2.4.	איכות השירות.....	12
2.5.	בניית אסטרטגיה במערכות שירות טלפוניות.....	13
2.6.	תמהיל המיומנויות הדרושות לניהול מוקדי שירות טלפוניים.....	13
2.7.	הטכנולוגייה התומכת.....	14
3.	ניתוב מבוסס מיומנויות – מהו?.....	16
3.1.	הקדמה.....	17
3.2.	המעבר מ-ACD ל-SBR.....	17
3.3.	תפקיד ה-SBR ב-CRM.....	18
3.4.	שימוש במאגרי מידע, כשימוש יעיל ב-SBR.....	19
3.5.	טכנולוגיה.....	20
3.6.	מורכבות התיכון וצעדים בהטמעת SBR – נקודת מבט תוך אירגונית.....	21
3.7.	SBR – האם זה תרופה?.....	21
3.8.	התאמת המיומנות לסוג ההתקשרות.....	22
3.9.	מיומנות יתר לעומת מיומנות לקויה.....	23
3.10.	מדידה ותרומת המיומנות.....	23
3.11.	ניתוח מבוסס מדיה.....	24
3.12.	מדיניות הכשרה לקידום – תגמול והכרה.....	24
3.13.	סיכום.....	25
4.	סקירה ספרותית.....	26
4.1.	Telephone Call Center: Tutorial, Review and Research Prospects.....	27
4.2.	An Introduction to Skills-Based Routing and its Operational Complexities.....	29

4.3. Resouce Pooling And Staffing Call Centers With Skill Based Routing.....	39
5. ניתוב מבוסס מיומנות – סקירת המתרחש בפועל.....	51
5.1. בנק פועלים.....	52
5.2. פלאפון.....	74
6. ניתוב מבוסס מיומנות – ניתוח מודלים.....	81
6.1. מודל V.....	82
6.1.1. הגדרת המודל והתסריט הנבדק.....	82
6.1.2. הגדרת הסימולציה.....	83
6.1.3. תוצאות הסימולציה וניתוחן.....	88
6.2. מודל N.....	103
6.2.1. הגדרת המודל והתסריט הנבדק.....	103
6.2.2. הגדרת הסימולציה.....	106
6.2.3. תוצאות הסימולציה.....	109
6.2.4. ניתוח התוצאות.....	112
6.2.5. הצעות למחקר המשך.....	114
7. סיכום.....	116
8. בביליוגרפיה.....	118
9. נספחים.....	119
9.1. תרשים הארגוני של "פועלים ישיר".....	120
9.2. תיאוריית ה-QED.....	121
9.3. קוד עבור מודל V.....	122
9.4. קוד עבור מודל N.....	128

פרק 1

מבוא

דו"ח זה מסכם את הפרויקט השנתי שערכנו בנושא ניתוב מבוסס מיומנות במוקדי שירות טלפונים. בשנים האחרונות ביסס את מעמדו נושא השירות, וזאת עקב מעבר מחברות יצרניות לחברות שירותיות (כ-80% מהכלכלה בעולם המערבי). כיום הבידול בין ארגונים שונים מתבסס יותר ויותר על השירותים הניתנים ללקוח ולא על המוצר. לשם כך הוקמו ברבות השנים מוקדי שירות רבים שתפקידם הוא יצירת קשר עם הלקוח, לפני ואחרי הקנייה. עקב עליית חשיבות השירות עלתה גם חשיבותם של מוקדי שירות הלקוחות.

מרביתם של מוקדי השירות הם מוקדים טלפוניים. המוקדים הטלפוניים מושתתים על מספר אבני בניין כפי שיוסבר בפרק 2 – מוקדי שירות טלפוניים מהם, ובמאמר המוצג בפרק 4.1. לאחר ביסוס מעמדם של מוקדי השירות הטלפוניים, נוצר מצב כי לקוחות שונים מבקשים שירותים שונים, וכן עם התפתחות תורת ה-CRM איבחנו כי לא כל הלקוחות זהים ובעלי ערך שווה לארגון. מכאן נוצר הצורך בהתאמת השירות הניתן במוקדי השירות לדרישות השונות של הלקוחות וכן ללקוחות השונים ע"פ ערכם. מכאן הגיע מושג ה"ניתוב מבוסס מיומנות" בו עוסק פרויקט זה. הסבר לגבי מהות ה-SBR (ניתוב מבוסס מיומנות, Skill Based Routing) מובא בפרק 3, ובנוסף יוצגו שני מאמרים בנושא.

מוקדי שירות רבים בארץ, עובדים כיום לפי הרעיונות של ניתוב מבוסס מיומנות. על מנת לחבר בי התיאוריה למציאות, במהלך הפרויקט ביקרנו בשני מוקדים כאלו – מוקד "פועלים ישיר" של בנק הפועלים ומוקד שירות הלקוחות של חברת הסלולאר "פלאפון". אודות חברות אלו, השירותים הניתנים בהם ובפרט ניתוב מבוסס מיומנות כפי שמובא לידי ביטוי בחברות אלו ניתן לקרוא בפרק 5.

לסיום בחנו שני מודלים שהם אבני הבניין של ניתוב מבוסס מיומנות – מודל V ומודל N. שני המודלים וניתוחם מוצגים מאגר אחד של שרתים ושני סוגי לקוחות, אשר לכל אחד מהם תור. הנושא שנבדק בנוגע למודל זה הוא שימוש במדיניות סף, על מנת להעלות את רמת השירות של לקוחות ה-VIP. המודל בודק מהי ההתנהגות של אותו סף אופטימלי עבור מקרים שונים, וכן מתי השימוש בסף זה הוא כדאי משיקולים של עלות. ההתנהגויות השונות של המודל נבדקו ע"י סימולציה מתאימה.

מודל N הוא למעשה צירוף של שני מודלים פשוטים מסוג M/M/N כאשר אחד ממאגרי השרתים יכול "לעזור" לשני על ידי טיפול בלקוחות של השרת השני במקרה של עומס. במחקרים קודמים נמצא כי מודל זה בעייתי מבחינת היציבות. פרויקט זה הוא המשך ישיר של המאמר: "An

Operational and its Introduction to Skills-Based Routing" (מנדלבאום וגרנט, 2003) אשר מציג את הבעייתיות האמורה במצב של שרת אחד בכל מאגר. הפרויקט בודק בעזרת סימולציה את ההשלכות של הגדלת מספר השרתים על היציבות של המערכת המדוברת.

פרק 2

מוקדי שירות טלפוניים,

מהם?

2.1 מוקדי שירות טלפוניים – מהם?

מוקדי שירות טלפוניים הם תוצאה ישירה של דרך החיים המתפתחת במאה העשרים. הצרכן הממוצע עובד יותר שעות, וכתוצאה מכך קטן הזמן הפנוי שברשותו. ככל שיורד שיעור הזמן הפנוי של הצרכן, כל המתנה או טרחה נתפסת כארוכה ומטרידה יותר.

במקביל לשינויים בדפוסי החיים של הצרכן בעשור האחרון, ועם התפתחות טכנולוגית התקשורת, הפכה מערכת הטלפון לכלי תקשורת נוח ויעיל. המערכת תומכת בשמירת קשר, העברת מידע וביצוע עסקים תוך התחשבות בשיקולי זמן ועלות- תועלת בביצוע עסקים ומתן שירותים.

שילוב התופעות המתוארות הביא להולדת תעשייה חדשה, היא תעשיית מוקדי השירות הטלפוניים (ה-Call Centers). עם הגידול בתעשיית מוקדי השירות הטלפוניים, התפתחו שיטות שיווק,

אסטרטגיות ניהול וטכנולוגיות נלוות.

הסיבה העיקרית להקמת מוקדי שירות טלפוניים ע"י חברות עסקיות היא יצירת ערוץ תקשורת זמין ונוח עם לקוחותיהם. הפרעות או עיכובים בתהליך יצירת הקשר עם הלקוח, פוגע ברווחיות החברה והצלחתה, ומוריד מערכה העסקי בעיני הלקוח.

ישנם גורמים רבים המשפיעים על תהליך יצירת הקשר עם המתקשר כמו: איוש המוקד, אורך השיחה, מספר הקווים, שעות ההפעלה וסבלנות הלקוחות. היות ולצרכנים יש פחות ופחות זמן לקניות ושירותים, הם יפנו למתחרים במקרה והם מרגישים שמופגנת כלפיהם אדישות או ניתן להם שירות גרוע.

חוסר הסבלנות של הצרכן, יחד עם הרגישות הגוברת של חברות לצרכי לקוחותיהם וחתירתם לרווחים, מביאים לכך שמנהלי מוקדי שירות טלפוניים מחויבים להיות במעקב מתמיד אחר טיב השירות הטלפוני, וליצור תהליך מתמשך של שיפור. על מנהלים לשלב מיומנויות רבות שיפורטו בהמשך, כגון - חקר ביצועים, סטטיסטיקה, פסיכולוגיה, כלכלה, תורת המשחקים, שיווק וחקר עבודה ושיטות. שילוב מיומנויות זה מביא לניהול אפקטיבי ויעיל יותר של מוקד השירות הטלפוני.

אבני הבניין של המוקד הטלפוני -

מוקד שירות טלפוני מושתת על ארבע אבני בניין מרכזיות:

- קווי תקשורת, המובילים שיחות טלפון ליעד מסוים.
- מכשירי מיתוג, הממיינים ומנתבים שיחות באותו יעד.
- סוכנים המקבלים את שיחות הטלפון או יוזמים שיחות יוצאות.
- מערכת מידע תומכת, בה נעזרים הסוכנים לאספקת מידע.

אפיון מוקדי השירות הטלפוני -

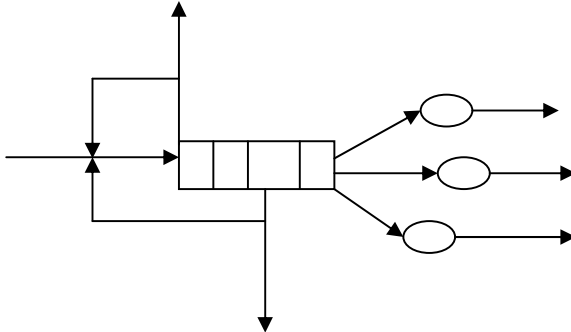
ניתן לחלק את המוקדים לארבעה סוגים עיקריים:

- מוקדי חירום, כגון משטרה, מד"א ומכבי אש
- מוקדי אינפורמציה, הנותנים או מקבלים אינפורמציה.
- מוקדי עסקים, המתעסקים במכירות, טלמרקטינג, מכירות דרך קטלוגים, שיווק ובנקאות.
- מוקדי שירות לקוחות ותמיכה טכנית.

חלוקה זו אינה תמיד חד משמעית, וייתכנו שילובים בין הסוגים השונים. מוקדי שירות טלפוניים מוגדרים בד"כ כשיחות נכנסות (inbound), כלומר, מקבלים שיחות, או כשיחות יוצאות (outbound), כלומר, מיועדים ליזום שיחות החוצה.

2.2 מידול מוקד שירות -

את המוקד הטלפוני הפשוט ביותר ניתן להציג במודל הבא :



- מודל זה מייצג מערכת תור המוכרת כ- $G/G/s/b$ - מופעים מהתפלגות כללית, זמן שירות מהתפלגות כללית, ו- s מוקדנים הנותנים שירות בסוף תור בעל b מקומות.
- במוקדי שירות טלפוניים בד"כ מניחים מופעים אקראיים לחלוטין ותהליך מופע פואסוני.
- שיחות מגיעות למערכת ע"י קווי טלפון חיצוניים למרכזיית המוקד. למוקד עצמו מספר מוגבל של קווים פנימיים לכל מספר טלפון הפונה אליו. מספר הקווים הפנימיים הדרוש תלוי בקצב הופעת השיחות, אורך השיחות ובמספר המוקדמים הנותנים את השירות, והוא שקבוע את אורך התור.
- שיחות המגיעות למערכת כאשר כל הקווים תפוסים יקבלו צליל תפוס וייתכן שינסו לבצע ניסיון חוזר להתקשר למערכת.
- השיחות המופיעות למערכת נקלטות ע"י ACD (Automatic Call Distributer), המנתב את השיחות למוקדן פנוי, באם קיים.

לעומת המערכת הפשוטה, קיימות גם רשתות שירות טלפוניות. ברשתות אלו מתבצעות העברות של שיחות בין מוקדי שירות שונים המפוזרים גיאוגרפית.

רשת אופיינית אחת היא רשת חזית ועורף. החזית זהה למודל המוקד הטלפוני הפשוט. מוקדני החזית מאומנים במגוון מסוים של שירותים, בהם הם מטפלים בלעדית. בנוסף קיימים גם מוקדני עורף, שיכולים להתפצל לפי מומחיות מסוימות, ואשר יכולים לטפל בבעיות ספציפיות הדורשות טיפול ממוקד יותר, או לחילופין משך זמן טיפול ארוך יותר.

2.3 מדדים לבחינת רמת השירות -

קיימים מדדים רבים לרמת שירות של מוקד שירות טלפוני, ואלה משתנים ממוקד למוקד. נסקור את המדדים הגלובליים לפי כרונולוגית התהליך:

- פרופורציית הלקוחות המתקשרים למערכת ומקבלים צליל תפוס.
- פרופורציית הלקוחות הנכנסים לתור ונוטשים בטרם קיבלו שירות.
- זמן מענה ממוצע.
- זמן המתנה ממוצע עד למתן שירות או עד לנטישה, בהינתן שהייתה המתנה.
- פרופורציית הלקוחות שלא המתינו כלל והועברו ישירות לשירות.
- אחוז הלקוחות הממתינים מעבר לפרק זמן נתון ולא נטשו את המערכת.
- אחוז הלקוחות הממתינים מעבר לזמן נתון ונטשו את המערכת.

מדדים שונים מבין הנ"ל חשובים יותר או פחות לפי סוג השירות אותו מספק המוקד. הרזולוציות בהם נבחן מדדים אלו משתנות, כך שניתן להסתכל על המדדים ברזולוציה חודשית, יומית ושעתית. או לחילופין להסתכל על זמני רגיעה וזמני שיא.

2.4 איכות השירות -

איכות השירות של מוקד תלויה לא רק בכך שהמוקד מציע מוצר או שירות המתאים לצרכי הלקוח, ומספקו ללקוח בצורה יעילה, אלא גם ביצירת אווירה ותחושה מספקת. איכות השירות קשה יותר למדידה ובקרה מאשר איכות מוצר המיוצר פיזית. השירות הינו חסר מוחשיות, הטרוגני וסימולטני. חוסר מוחשיות משמעותו כי קשה למדוד את מידת עמידת השירות בדרישות הסף, משום שהאלמנט הפסיכולוגי של השירות קשה למדידה ולצפייה. ההטרוגניות נובעת מכך שנותני השירות והלקוחות הם חלק קריטי של השירות, והם בעלי אישיות שונות ומגוונות שאינם ניתנות להכללה לתבנית או לבקרה. הסימולטניות של השירות היא בכך שבהסתכלות על השירות כמוצר, ייצור השירות והשימוש בו מתבצעים בו זמנית, כך שלא ניתן לבדוק או לבקר את מוצר השירות לפני שהוא נמכר ללקוח.

2.5 בניית אסטרטגיה ניהולית במערכת שירות טלפונית -

ניתן לחלק את האסטרטגיה הניהולית לארבעה תחומים: תכן, בקרה, תפעול, וגורמי אנוש.

תכן - קיימות מספר בחירות שעל מתכן המערכת לעשות:

1. בחירה בין מערכת גמישה, בה כל השיחות מגיעות לתור אחד, לבין מערכת מומחית, בה לכל סוג שירות תור ומוקדן משלו.
 2. הבחירה בין חזית לעורף, כאשר השימוש בשירותי העורף יקר יותר ושימוש בהם מאריך את זמן התגובה.
 3. בחירת האיזון בין נטישות לצליל תפוס, כאשר הגדלת מספר הקווים בהכרח תגרום להמתנה רבה יותר לעומת הקטנת מספר הקווים שתגרום להורדת זמן ההמתנה הממוצע של הממתנים בתור והגדלת פרופורציות המתקשרים המקבלים צליל תפוס.
- בקרה - מדידות שוטפות הנם תנאי הכרחי לניהול נכון של המוקד השירות הטלפוני.
- תפעול - לאחר אפיון המופעים, הנטישות ותהליך השירות של המערכת, יש לקבוע את רמת האיוש של מוקד השירות לכל יחידת זמן ואת מספר הקווים הדרושים בכל יחידת זמן.
- גורמים אנושיים - אמידת מידת הסבלנות של האוכלוסייה והערכתה, הערכת פרופורציית הנוטשים המתקשרים חזרה למערכת ותפיסת הזמן שהלקוח מצפה להמתין בתור עד שיענו לו.

2.6 תמהיל המיומנויות הדרושות לניהול מוקדי שירות טלפוניים -

א. חקר ביצועים -

- תחום זה מקנה את רוב המדדים והכלים המתמטיים לתכן נכון של מוקד טלפוני, ולפתרון בעיות הנובעות מתכן לא נכון. התחומים הנכללים במיומנות זו הינם:
1. מידול המערכת בעזרת מודלים של מערכות תורים ורשתות תורים.
 2. עיצוב ארכיטקטורת המערכת.
 3. קביעת מספר המוקדנים.
 4. קביעת מספר הקווים.

- ב. פסיכולוגיה -
הפרמטרים שצריך לקבוע הנם: קביעה של גישת השירות, הערכת סובלנות האוכלוסייה, השפעת ממד הזמן הנתפס בעיני הלקוח, הערכת מידת החיובים החוזרים ושחיקת העובדים.
- ג. כלכלה -
על מנת לבדוק את הכדאיות הכלכלית של המוקד והרווח הרצוי נשתמש בשני פרמטרים לבדיקה: האחד עקומות עלות תועלת והשנייה איזון בין תפוקה לעלות בעקבות שיפור השירות.
- ד. תורת המשחקים -
זהו כלי למציאת שיווי משקל כלכלי בין עלות לתועלת וכן לניבוי קצב הנטישה.
- ה. מערכות מידע -
מוקדי שירות טלפוניים נתמכים ע"י בסיס מידע ומערכת מחשב. מערכות אלו צריכות לתמוך בתהליך השירות ולקשר למערכות מידע מקומיות ועולמיות.
- ו. שיווק -
מטרת השיווק הינו להרחיב את מעגל הלקוחות. לכן, יש לקבל 2 החלטות מרכזיות: האחת הנה מיצוב המוקד והשירות הניתן והשנייה הנה משיכת הלקוחות החדשים למערכת.
- ז. סטטיסטיקה -
על מנת לקבל החלטות נכונות אנו מתבססים על נתונים סטטיסטיים רבים. על סמך נתונים סטטיסטיים אנו מנתחים את הנתונים, בונים פונקצית תועלת וכן בונים תחזיות.
- ח. חקר עבודה ושיטות -
במסגרת זו נקבעים נהלי העבודה הנכונים. ע"י המדידות אנו מנסים לבנות מודל אנליטי למוקד. כמו כן, בעזרת חקר שיטות עבודה אנו מנסים לקבוע שיטות עבודה מיטביות, לחקור את התנאים הפיזיים של העובד במטרה לשפרם וכן לקבוע שיטות לשכר עידוד.

2.7 הטכנולוגיה התומכת -

השירות הטלפוני הוא עתיר טכנולוגיה. משולבות בו מערכות תקשורת מורכבות, מערכות מידע ובסיסי נתונים גדולים, כלי אחזור מתקדמים ותוכנות מתקדמות לאיסוף מידע ומדדים. כל מוקד שירות טלפוני נתמך ע"י תשתית טכנולוגית. מעבר לצרכים בסיסיים אלו, התשתית הטכנולוגית משמשת גם כאמצעי לאיסוף נתונים לגבי הקורה במערכת, מהם נוכל להפיק את המדדים הנחוצים לקביעת רמת השירות.

את הטכנולוגיה הקיימת ניתן לחלק לשלושה סוגים:

(1) מערכת מיתוג מסוג ACD - מערכות אלו מקבלות את השיחות הנכנסות מקווי טלפון ומנתבות

אותן לעמדות הטלפונים של המוקד. מערכות ACD בד"כ אוספות נתונים בסיסיים בלבד על המערכת, ואלו מוצגים רק ע"י אחוזים וממוצעים.

(2) מערכות איסוף מידע - מערכות אלו נולדו בעקבות הצורך לנהל מעקב אחר שיחות יוצאות. המידע שנשמר הינו משך השיחה, הקו ממנו יצאה השיחה וכו'.

(3) מערכות משולבות מחשב - מערכות ה- CTI (computer -telephone integration) הן

המערכות המתקדמות ביותר הקיימות בשוק התשתיות של מוקדי שירות טלפוניים. ניתוב השיחות, איסוף המידע וחישוב המדדים נעשים כולם ע"י תוכנת מחשב. מערכות אלה מאפשרות גמישות רבה למנהל המוקד, בעקבות האפשרויות הרבות שהן מציעות לניהול שיחות, ובשל כמות הנתונים המתקבלים ורמת הפירוט שלהם. גמישות מערכת זו מאפשרת להנהלת המוקד להציב מוקדנים מומחים בתחומים שונים בחזית ומאחוריהם להציב את מוקדני העורף המומחים.

ניתן לחלק את המערכות הטלפוניות מבחינה גיאוגרפית ל- 2 מערכות -

(1) מערכות Distributed Call Centers - אלו מערכות משולבות מחשב, המאפשרות איחוד

מוקדים באיזורים ג"ג שונים למערכת מקושרת. חלוקת השיחות בין המוקדים הג"ג השונים יכולה להיעשות עפ"י האזור ממנו הגיעה השיחה או בעזרת גלישה ממוקד למוקד.

(2) מערכת Virtual Call Centers - מוקדים אלו מורכבים ממספר רב של עובדים, אשר מעניקים

את השירות הטלפוני מביתם, אך מקושרים למערכת מרכזית המבצעת את כל הפעולות שמבצעת מערכת מרכזית במוקד טלפוני פיזי.

פרק 3

ניתוב מבוסס מיומנות, מהו?

3.1 הקדמה -

במוקדי שירות טלפוני ניתן להגדיר את מדיניות ה- SBR (Skill-Based Routing) כיכולת לנתב שיחה אל הנציג בעל היכולת הגבוהה ביותר לענות לשיחה זו.

SBR לוקח בחשבון את דרישות הלקוח הנובעות מסיבות שונות – החל מהבדלי שפה ועד הצורך באינפורמציה ספציפית לגבי מוצר מסוים. השיטה יכולה להעריך את ערכו של לקוח על פי פעולות העבר שלו ולשער באמצעותן את התנהגותו בעתיד. זה אולי נשמע קל לביצוע כאשר חושבים על זה כעל העברה פשוטה של שיחה אל הנציג המוסמך, אולם האתגר שבמולטימדיה ובמוקדי התקשרות הפועלים באמצעות האינטרנט איננו פשוט כלל ועיקר. SBR בעצם מספקת הזדמנות טובה למקסם את הפוטנציאל של מרכז השירות- במידה וידוע כיצד יש לתפעל זאת לטובת המערכת הנידונה.

3.2 המעבר מ ACD (automatic call distribution) ל SBR -

עם ההתפתחות הטכנולוגית, מתפתחת הדרך בה מנוהלים מוקדים טלפוניים. ההתקדמות מהשיטות הישנות, האוטומטיות אשר ניתבו שיחות ע"פ פרמטרים סטטיים, לניתוב מבוסס המיומנויות הנוכחי מצביעה על השינוי המהותי המתרחש בתפיסת המוקדים הטלפוניים.

מודל ה-ACD הסתמך על הנצילות, או כמות האנשים בתור ודאג לניתוב שיחות לנציגים השונים במהירות האפשרית. ניתוב השיחות נקבע על פי פרמטרים ידועים, אשר היו מבוססים על חוקים סטטיים שלא התאימו להגיון העסקי של הארגון. מכל הנאמר לעיל לא היה יחס ישר בין הצורך של הלקוח לבין מיומנות הנציג שטיפל בשיחה שלו. נהלי ה ACD התמקדו בעיקר בהפחתת העלויות באמצעות מתן מענה לשיחות במהירות הגבוהה ביותר האפשרית.

בעבר, כונו מוקדים טלפוניים מרכזי עלות. כיום המצב שונה - חברות רואות את הפוטנציאל הטמון בהפיכת המוקד הטלפוני שלהן למרכזי רווח. מרכזים טלפוניים כיום משמשים יותר כאלמנטים אסטרטגיים בניהול העסק. אסטרטגיית SBR עוזרת להפוך זאת לאפשרי באמצעות יצירת יחס ישיר בין צרכי הלקוחות וכישורי הנציגים. בזמן ש ACD שימשה בעיקר כדי להפחית עלויות, SBR יכולה לתרום ביצירת רווח. ACD ניתבה שיחות לכל נציג פנוי או לכל תור באופן כזה או אחר עד אשר נציג התפנה. SBR לעומת זאת מנתבת כל שיחה אל הנציג הספציפי שיכול להעניק את השירות הטוב ביותר ללקוח ובאופן משמעותי יכול להפחית את זמן ההמתנה של שיחות שלא ניתן לתת להן מענה באופן מיידי.

שיטת ה SBR גם הרוויחה מהחשיפה לטכנולוגיות חדשות. בעבר זיהוי הלקוחות נעשה בדרך כלל על פי מספר שהלקוח התבקש לחייג או על פי המספר ממנו הם התקשרו למוקד. כיום, מערכת ה-IVR (interactive voice response) יכולה לבקש מהלקוחות להזדהות. עם האינפורמציה הזו, המערכת יכולה לבדוק את בסיס הנתונים ולהשתמש בתכולת התיק של הלקוח כדי לנתב את השיחה. ניתוב השיחות מבוסס על הפרופילים של הלקוחות וההיסטוריה של שיחות הלקוח. מכאן, הסיבות השונות שבגללן יצר הלקוח קשר עם המוקד בעבר ותמרון ה-IVR, מאפשרים לארגון להיות מתואם עם הצרכים והדרישות של הלקוח ועל ידי כך להעלות את רמת השירות על ידי קיצור זמני ההמתנה, ולהשתמש ביכולות סטטיסטיות בזמן אמת כדי לקבוע את ניתוב השיחה. כל אלו אף יכולים להביא לגידול ברווחי הארגון.

3.3 תפקיד ה- SBR ב- CRM (customer relationship management) -

במידה ו- SBR היא תוצאה של ההתפתחות הטכנולוגית, במידה שווה היא נוצרה ככלי עבור פילוסופיית ה- CRM לפיה יש לפתח ולשמר את היחסים עם הלקוחות, בהתבסס על ערכם עבור הארגון.

מאחר ו- CRM נוגעת בתהליכים, באנשים ובטכנולוגיה, אסטרטגיה מונחת CRM אמורה לכלול את הצורך בזיהוי לקוחות וערכם, על מנת להגדיר כיצד יש לשרתם וכדי להכשיר עובדים לטיפול ביחסים עם הלקוחות. SBR הפכה לכלי המאפשר שימוש ב- CRM במרכז הטלפוני. כאשר מוצרים, הצעות שירות וכד' דומים, המוטיב המשפיע הופך להיות טיב השירות הניתן. קצב הליקויים בתהליך שירות לקוחות קשור באופן ישיר לטיב הטיפול בלקוחות, ולקוחות יעברו לכל מי שרק יציע את השירות הטוב ביותר. טיב השירות לכן הופך להיות למוטיב קריטי וחשוב ביותר. בשנת 2000, אוניברסיטת Purdue סיכמה ש 68% מהלקוחות שנותרו נאמנים לארגון עשו זאת בגלל הטיפול הטוב שהם קיבלו. מאז עדכנה האוניברסיטה כמות זו ל 78%. מידע זה אפילו יותר רלוונטי, כששוקלים את העובדה שעלות השגת לקוח חדש לארגון גדולה פי 4 עד 11 משמירה על לקוח קיים. לכן יש סיבה טובה לפעול לשימור לקוחות ופיתוח היחסים עמם מבחינת התועלות לארגון ומבחינת ההחזר על ההשקעה- דבר המשפיע באופן חיובי על השורה התחתונה.

האם כל הלקוחות שווים?

האם זהו שיקול עסקי נכון, לשמר את כל הלקוחות?
התשובה הפשוטה היא כן. אולם זהו כן מסויג, מפני שעלות הטיפול בחלק מהלקוחות גבוהה מההכנסה שהם יוצרים. ארגונים לא מעוניינים "להיפטר" מלקוחות, לכן נכנס האלמנט המרכזי של אסטרטגיית ה- CRM - סיווג לקוחות. כאשר לקוחות מחולקים לקטגוריות ניתן לאפיין את הטיפול בכל קטגוריה.

לדוגמה - לקוחות פחות רווחיים יופנו יותר לשירות עצמי ואילו לקוחות רווחיים מאוד יקבלו עדיפות עליונה, כך שהארגון יכול לבנות ולשמר אתם יחסים ארוכי טווח. כאשר מיישמים מיון לקוחות, יש לזכור לזהות לקוחות בעלי פוטנציאל להפוך ללקוחות רווחיים יותר. במילים אחרות יש לזהות "כוכבים עולים". אזי הפקטורים הנלקחים בחשבון הם לא רק הקיימים אלא גם הפוטנציאלים.

מערכת ה- SBR יכולה להשתמש בפרופיל העדפות פרטני אף יותר. היא יכולה להסתמך על היחס 80-20 של רמת שירות המבוסס יחסית אך עיוור במידה מסוימת. יחס זה הוא יקר וקשה לשמרו לאורך זמן.

לכן פעילות עסקית חשובה נוספת יש לשקול כאשר מתכננים את מערכת ההעדפות עליה מבוססת ה- SBR בארגון והיא ערך הטרנזאקציה. במידה ולקוח מתקשר כדי לשלם חוב מאחר, שעולה בזמן, אנרגיה וכסף כדי לספק את הטיפול, או שלקוח מתקשר כדי לבצע רכישה כלשהי, טראנזאקציות אלה צריכות לקבל עדיפות גבוהה יותר מטרנזאקציות אחרות כמו שינוי כתובת שהינה חסרת ערך פוטנציאלי רווחי. את אופי התעדוף ניתן ליישם לכל סוג של קשר, טראנזאקציה או לקוח. לכן הארגון חייב לשאול את עצמו, מדוע רמת השירות הנוכחית שלו היא כמו שהיא וכיצד ניתן לשפרה כדי להגדיל את הרווח.

3.4 שימוש במאגרי המידע כשימוש יעיל ב- SBR

כפי שניתן לראות, SBR היא הרבה יותר מסתם שיטה לניתוב שיחות לשרתים מיומנים. למעשה, SBR משלבת בין אופני ניתוב שונים המבוססים על נתוני העבר מניתוב מונחה דרישות לקוח – מציאת נציג מתאים, ועד ניתוב מונחה לקוח- סיווג רמת ההכנסה של הארגון ממנו.
כל זאת מצביע על יכולתה של ה- SBR לאמץ מגוון רחב של תכנים עסקיים, אסטרטגיות וטכנולוגיות. המונח- "ניתוב מבוסס מידע" מצביע על החשיבות שבגישה והיכולת אגירה של אינפורמציה חדשה.
אחד המפתחות להצלחה הוא מיפוי נכון של המידע והפיכתו למידע עסקי המכריע החלטות עסקיות. כדי להטמיע אסטרטגיית SBR, יש להתחיל בקביעת כללים המושתתים על תהליכים עסקיים מושכלים, המבוססים על האופי השיווקי של הארגון, המכירות ומאגר המידע של המוקד.

בשלב זה ניתן להשתמש בכללים כדי לפתח תסריטי SBR, אשר יכולים ליצור מידע עסקי מוצק העוזר בתיכון ועיצוב האסטרטגיה ויכול לתרום לשיפורה.

עם שימוש נכון, SBR יכולה לספק מידע עסקי עשיר ושימושי שיכול להאיר תחומים עם פוטנציאל לרווח כמו אלו שהוזכרו – ולזהות תחומים רגישים לפני שאלה הופכים לבעיות. מנהלים יכולים להתמקד בניהול משאבים והשגת מטרות הארגון דרך איסוף מידע ודיווח, יותר מאשר דרך ניתוח תנועת הנציגים מתור אחד לאחר והתפלגות עומסי השיחות המגיעות. SBR יכולה גם לספק דוחות המפקחים על הפעילות אל מול הלקוחות ומכילים אינפורמציה רבה על פעילות המוקד. מנהלי מוקדים הופכים עתה לאנשים חשובים במובנים האסטרטגיים של הארגון, שימוש בדוחות ה - SBR כדי לזהות מגמות והזדמנויות במכירות. במוקד הטלפוני הנציג הינו הקול של החברה. ומנהל המוקד הפך להיות המפקח האסטרטגי שלו. בגלל הכמות הפוטנציאלית של המידע, יש צורך לדעת כיצד ללכוד את התמונה הרצויה עוד בטרם מיישמים את השיטה. עניין זה מדגיש את החשיבות של תיכון יעיל כדי להבטיח שאסטרטגיית הניתוב אכן מפיקה את כל הסטטיסטיקה הנחוצה להתפתחויות בעתיד. מערכת הדיווח לכן הופכת לתת מערכת, והדרישות הטכניות חייבות להילקח בחשבון, יחד עם הדרישות העסקיות המשמשות בסיס לפונקציונאליות של ה - SBR. כאן יכול הארגון בעצם לתעל עצמו במנגנון המספק דרך התמקדות עסקית חכמה יותר.

3.5 טכנולוגיה -

מבחינה טכנולוגית, SBR הוא פתרון וירטואלי. לדוגמא הקשר בין פיתרון ה- SBR לבין שרת האינטרנט של הארגון איננו בהכרח פיזי. אותו פתרון SBR ניתן ליישום במספר מיקומים שונים בארגון, מאחד את המוקד הטלפוני למוקד וירטואלי הנתמך ממקום אחד. כל שרת מיומן פנוי יכול לענות כאמור לכל פניה, ללא תלות במיקום המקורי אליו הגיע השיחה. או במיקום הנוכחי של השרת. יכולת זו יכולה בעתיד אף לתמוך בשרתים הפועלים מהבית, אלטרנטיבה אטרקטיבית בשוק תחרותי. ה- SBR מציג בנוסף את התור הווירטואלי, ההופך את המוקד הטלפוני למוקד תקשורת. התור הווירטואלי מבטיח שאותם כללי ניתוב, כללים עסקיים, אסטרטגיית CRM ויכולת הדיווח יאפשרו לטפל בכל סוג של אינטראקציה; לכן המוקד מייצג גוף שלם, עם קול אחד המייצג את הארגון מנקודת המבט של הלקוח.

3.6 מורכבות התיכון וצעדים בהטמעת SBR - נקודת מבט תוך ארגונית -

הטמעת ה-SBR יכולה להוות מעין *הלם סיסמי ראשוני* לרוב החברות הקיימות. שיטות ההעסקה הקודמות, המבוססות על ותק או הערכת יעילות הביצועים מתערערות משום שהשיטה החדשה משפיע על איוש המרכז. מלכתחילה SBR לכאורה נראית כבעיית ניתוב פשוטה – עבודה עם אלגוריתם פשוט המבוסס על כישורים זמינים למען יצירת מנגנון אספקה מתאים ללקוחות. אולם חברות, אשר הטמיעו SBR גילו שהזעזועים וקשיי ההסתגלות הנלווים חזקים וארוכי טווח.

3.7 SBR – האם זה תרופה?

השינוי הראשון שנדרש הינו שינוי בגישה. בעבר שיחות חולקו למחלקות מסוימות, למשל, שאלות חשבונאיות, דווח שגיאות או הדרכה במכירות, ולמחלקות אלה הוקצבו צוותים מסוימים לטיפול. בדרך זו אפשר היה לשבץ את כל השיחות הנכנסות של הארגון לאחת הקבוצות, אשר אומנה ועל בסיס הניסיון הגיע לרמת מיומנות גבוהה מאוד של הנציגים. יחד עם ההתקדמות והגדלת הכושר בתוך המחלקה, הידע הגובר לא הגיע לנציגים מחוץ למחלקה הספציפית, ולכן הידע והניסיון הנצבר לא הועברו דרך הארגון כולו. הבלתי נמנע הוא שהלקוח ה"מורכב" שצריך, לדוגמה, דיון כלשהו, יטופל ע"י נציגים מצוות טכני מיומן אולם עליה בביקוש לנציגים מתחום התמחות זה, המוגבל בעיקרו, יגרום להגדלה בזמני ההמתנה ספציפית למחלקה זו.

בכל צעד שנעשה להקלת הקושי הזה נמצאו שיחות המגיעות לצוותים שלא מיומנים לטיפול בסוג זה של לקוחות, מהלך שהביא לאי-שביעות רצונם של הלקוחות וריבוי שיחות החוזרות כדי לפתור מחדש את הבעיות שנפתרו חלקית או בצורה פשטנית.

הבעיה הזו בנוסף גרמה להגברת הלחץ על המשאבים או למספר גדל של העברות שיחות פנימיות, כאמור, הצוות שמקבל שיחה אשר לא מיומן לטפל בה, מחזיר אותה לתור הלקוחות המופנה לנציגים המקוריים. תהליך של ריבוי קבוצות שהנציגים בהן יכולים להשתייך גם לקבוצות התמחות נוספות נחשב לדרך טובה להתמודדות עם הבעיות שצוינו אולם שיטה זו נתקלה בהתנגדות הולכת וגוברת מצד ארגונים שהתקשו להתמודד עם הצורך ליצור קבוצות מתאימות של צוותים והתמחויות.

בסביבה כזו חברות שונות ערכו הכרות עם SBR והשיגו הצלחה מיידיית עם היכולת שלה להתמודד עם מצבי לחץ במרכז תקשורת מודרני. בסופו של דבר, נציג עם מיומנות רלוונטית יכול להיות משובץ ללא תלות בצוות אליו היה שייך קודם. יכולת לנתב את הלקוחות לנציגים המיומנים לטיפול ראשוני הסירה לחץ "מלאכותי" פנימי על הנציגים, הפחיתה כמות העברת שיחות פנימיות ולכן הפחיתה את העומס עליהם. כמובן, זה גרם להקטנת זמני המתנה וכתוצאה מכך, שיפור ברמת שביעות רצונם של הלקוחות, הגביר היקף הטיפול ורמת שירות, הוריד את מחירי השירות, ואם הטיפול מבוצע נכון, יכול להוביל לשביעות רצון גבוהה בקרב העובדים.

3.8 התאמת המיומנות לסוג ההתקשרות -

הצעד הראשון בתהליך הטמעת שיטת ה SBR הוא לזהות את המיומנויות הנכונות הדרושות לטיפול בסוג שיחה מסוים. אולם צעד זה אינו פשוט – יש צורך בהגדרה מחדש של הכישורים הזמינים ההתמחויות הקיימות שבד"כ אינם מתאימים למערכת החדשה. החלטות על אילו נציגים קיימים כדי לטפל בבעיה מסוימת והתאמת מיומנותם לסוג ההתקשרות מתחיל, מהסוף הלא נכון של הבעיה. הגישה הנכונה יותר היא גישה ה-"top-down", בה שיחה מוגדרת ללא קשר למיומנויות הקיימות בארגון. זה יכול להיראות שגוי או לא הגיוני, אך זהו האידאל החדש של השלמות, שמביאה איתה שיטת ה SBR כחלק מהשפעתה הראשונית על תרבות השירות.

ראשית מוגדרת השיחה במונחים של האלמנטים הנדרשים לפתרון בעיה או במתן מענה לדרישה, תוך הנחה שהנציג המושלם בעל המיומנויות הנדרשות כבר קיים. תהליך זה, לעיתים קרובות, חושף מיומנויות אשר טרם הוגדרו או נמדדו, לדוגמה, הכרת המוצר או כישורים במתן שירות. השיטות הרגילות של אימון והערכה יכולות אומנם למדוד בהצלחה את יכולות הנציג בטיפול בכמות לקוחות מסוימת או בסוג לקוחות מסוים, אך לעיתים רחוקות ניתן לגלות את יעילות השירות של אותו נציג. פעמים זה עלול להיות כישור בודד מרכזי שיהיה חשוב ביותר בתסריט מסוים. למשל, אם הלקוח מתקשר לארגון בענייני בעיה בחיוב חשבון – המיומנות הנדרשת, כמובן, היא ידע מפורט על המחירים ותהליך הגביה (מיומנות רגילה בחשבונאות, ואפילו שילוב אם ישנה מורכבות גבוהה!). אך אספקט שיכול להתפסס והוא בעל חשיבות קיצונית ללקוח הממורמר שנאלץ להתקשר לחברה במטרה לתקן שגיאה שאין הוא אחראי לה, הוא שהלקוח מעוניין לפגוש נציג שיכול להזדהות איתו, שמבין את דאגתו ויחד עם זאת יכול לפתור את הבעיה.

כאן זו המיומנות הסמויה, הנפרדת ממיומנויות אחרות ומוגדרת כיכולת ל- "שירות הלקוח", אשר מצביעה על הבדל בין בעיה שנפתרה לבין בעיה שנפתרה והלקוח מרוצה. הימנעות מחוסר שביעות רצון ממושך של הלקוח שנוצר בעקבות הטעות שנעשתה, יכולה להיות חשובה, כמו תיקון הטעות עצמו. בנקודה זו לשיטת ה SBR יש השפעה מיידית.

הדרישות ה "אידיאליות" הללו מאלצות את הארגון להצדיק את התהליכים הנוכחיים ומנגנון התמיכה הארגוני בלקוחות ובאינטרסים שלהם, עניין המהווה משימה קשה תמיד. לעיתים קרובות מתפתחת בארגון מעיין צמיחה טבעית, וההתנסות בהצלחה, או יותר מדויק, חוסר ההתנסות בכישלון משמעותי, מאפשרים את התיכון ה"טבעי" של התהליך.

ההלם הראשוני יכול להוביל אפילו לרציונליזציה או שינוי דגשים באימון העובדים ובשיטות הטיפול בשיחות המוביל למתודות אידיאליות המוגדרות לפי הפונקציונאליות הנדרשת לתמיכה בשיטת העבודה החדשה.

בדוגמה שהוצגה קודם, פתרון בעיה חשבונאית יכול עתה לכלול לא רק טיפול על ידי מחלקת החשבונות, אלא מיומנות גבוהה יותר בשיטות חיוב, הכרות עם המוצר, תהליך גביה ומתן שירות ללקוח.

הדרישה החדשה הזו של הפרקטיקה הקיימת יכולה להפתיע את הארגון אך גם ליצור מודעות למיומנויות שכבר קיימות אך טרם זוהו עד כה בקרב צוות הנציגים הקיים.

3.9 מיומנות יתר לעומת מיומנות לקויה -

לא ניתן להתעלם מהשפעת שיטת ה SBR על הארגון ובמיוחד על העובדים בו, למרות שהקשר בין זמני המתנה, אורכי התורים ועלות השירות לשיטה לא תמיד פשוט. בתעשייה, בה תחלופה השנתית של העובדים עומדת על 30-50% לשנה (תלוי במדינה, בתעשייה ובסוג ההתקשרות), תחזית הצופה את תחלופת כוח האדם הנוכחי תוך כשלוש-ארבע שנים אינה נאה במיוחד.

הגדלת השכר והוספת גמישות למשרות יכולים להאט קצב האובדן, אך עם התחרות בשכר הקיימת בתחומי התמחויות שונים, בשילוב עם אימון עובדים חדשים (עד שישה חודשים במספר תחומים טכניים), ערך הפקטורים הללו קטן. נציגים לומדים לשווק עצמם טוב יותר בין הארגונים המתחרים, חושפים את יכולותיהם הכוללים אלמנטים בסיסיים שהינם בעלי ביקוש בין ארגונים שונים הנותנים שירות.

SBR מובילה את הארגון להערכה מחודשת של השיטות הקיימות לטיפול בשיחות, לשדרוג תהליכי תמיכה, וזיהוי אלמנטים כמו ניסיון וכישורים בסוג המיומנות הנדרש. חברות גדולות הכניסו את הפקטור הזה אל תוך אלגוריתם זיהוי המיומנויות שלהם, דבר שנתן להם אפשרות להגיע להחלטות ניתוב יותר מורכבות ובניית מודלים יותר מדויקים, המייצגים את רמת המיומנות הנדרשת או קיימת בארגון.

הארגונים שנוהגים לעקוף את הרעיון העומד מאחורי ה SBR , כלומר מסווגים העובדים כוותיקים וצעירים וניתוב השיחות נעשה בהתאם לוותק בלבד, ללא התחשבות בכישורים הספציפיים, מסתכנים באובדן היתרונות האמיתיים של SBR ופשוט מקימים צוותי נציגים בעלי "נכות" לקליטת שיטות עבודה כמו ה - SBR. ההלם השני הזה בעצם גורם לתחילת הזעזועים בתחומים האחרים שלא התקשרו קודם ל - SBR.

3.10 מדידה ותרומת המיומנויות -

מרכזי התקשרות לעתים קרובות נראים בעלי חשיבות נחותה בהשוואה לחלקי העסק האחרים, אך בעזרת ה SBR יש להם ההזדמנות לבסס את עצמם ואת רמת המקצועיות שלהם בתוך הארגונים הגדולים. בעבר, שכר ותנאי העבודה היו תלויים במורכבות ההכשרה שעבר הנציג ובמספר העיסוקים השונים בתפקיד, כאשר כל אלה מיוחסים ישירות לוותק. SBR, בעזרת הגדרת תחומי הכשרה וביצוע ראיונות אישיים מספק לארגון לראשונה סווג אישי המוגדר היטב. הפרופילים הללו מאפשרים לקיים חלוקה על פי רמת הנכונות ועומק הידע, כישורים המשמשים מפתח בהערכת הנציג ובקידומו.

האלמנט החדש, פיצוי כספי של הנציגים, יכול לתרום בזיהוי ועידוד של מיומנויות באופן רחב יותר באמצעות המרכז, כאלה המתאימות לדרישות הדוחקות של עסקים, לדוגמה בתמיכה ביוזמה שיווקית חדשה. סוף סוף יוכל המנהל לקבל תמונה אחידה של המיומנויות העומדות לרשות הגוף שהוא אחראי לו. עובדה זו מאפשרת לקבל החלטות מושכלות יותר ויכולת להכריע בהחלטות לגבי תיכון תהליכים בהמשך וביצירת יוזמות חדשות.

ההלם השלישי מצייד את מנהל המרכז במידע הנחוץ לו כדי לייצג את המחלקה באופן מדויק יותר כחלק מהארגון הכולל, תוך יצירת הכרה רבה יותר בערך ועומק הידע שנצבר במחלקה שלו. תמיכה בהכרה זו היא היכולת להעריך יכולותיו של נציג וקביעת התגמול המתאים לנציג הספציפי על בסיס מיומנות מסוימת. באמצעות צעדים אלה מתגבש לו ארגון המבוסס על תהליכי למידה.

כעת ניתן לקבוע מסלולי קידום המבוססים על פרופיל כישורים, וכל נציג יהיה מצויד בתמונה ברורה של המיקום שלו במבנה המרכז ובאילו צעדים עליו לנקוט כדי להגדיל את המוניטין האישי שלו. ייפוי כוח שכזה הניתן לנציגים מעודד אותם להיות פעילים במרדף אחר הכשרה ולמידה תמידיים ולא כפופים יותר לנוהלי הכשרה קפואים בלבד.

3.11 ניתוב מבוסס מדיה -

עם ההתקדמות ממוקד שירות טלפוני למוקד מולטימדיה המספק שירות, נדרשים כישורים נוספים. ארגונים שונים אכן צפו את המגמה הזאת והציגו תורים אחידים על פני מדיות שונות לפיהם יש לבחור נציגי שירות.

אולם ניתוב המבוסס על סוגי מדיה מצריך כישורים המותאמים באופן ספציפי לסוג המדיה, לכן כאן נכנסת שוב החשיבות של "פרופיל מיומנויות אידיאלי". טיב הדרישה יכול להיבחן ופרופיל ידיאלי ניתן לפיתוח. לצורך כך יש להעריך את מבנה הבעיה של הלקוח שלא תמיד מתורגמת לשפה מפורשת אלא לסדרה של פעולות למשל. אולם ברגע שזו זוהתה היא יכולה לדרוש טיפול פשוט יחסית למשל טיפול בדואר אלקטרוני או בקשות יזומות דרך האינטרנט. היכולת לפרש את הבקשה ולזהות אולי הזדמנות לעסקה יכולה להניב מכירה שיתכן ושרת אחר, פחות מיומן לא היה מצליח ליזום.

כאשר מטפלים בבקשה מילולית של לקוח, זה קל יותר לזהות את מצב הרוח של הלקוח ע"פ נימת הדיבור, והטונים אולם זיהוי זה קשה יותר כאשר מדובר בבקשה כתובה כלשהי. לעיתים זה קל יותר לנציג לפתור אי הבנות בשיחת טלפון מאשר לנסות ולפתור את הבעיה בכתב.

כדאי להזכיר שכל מסמך המועבר בין הארגון ללקוח עלול לשמש כראיה כנגד הארגון מבחינה משפטית במקרים מסוימים. הגמישות שבנגישות הארגון ללקוח היא *ההלם הרביעי* שיוצרת ה SBR. זהו הצורך להתמודד עם מגוון רחב יותר של פרופילי כישורים של נציגים, החל ביכולת הפשוטה לתת מענה טלפוני וחלה בפתרונות מבוססי רשת המצריכים רמה מסוימת של ידע טכנולוגי.

3.12 מדיניות הכשרה לקידום - תגמול והכרה -

נציגים יכולים למקם עצמם בצורה יותר מדויקת בהיבט הרחב יותר של המוקד, וכעת יש להם האפשרות להחליט באיזו הכשרה הם היו מעוניינים כדי להעלות את רמת המיומנות שלהם, ולתת

מענה למגוון רחב יותר של בעיות המגיעות למוקד. בנוסף, מנקודת המבט של הנציג, הוא יכול להעריך כעת מה נחוץ מבחינתו כדי לקדם עצמו בתוך הארגון, ואף להיות מתוגמל טוב יותר. מנקודת המבט של הארגון, יש לארגון כעת את היכולת לנתב מגוון רחב יותר של סוגי התקשרות לנציג המתאים מבלי הצורך בתורים מורחבים או העברת פניות אך בנוסף עליו להציע מסלולי תגמול שילוו מערכת זו.

ניתן להציג סקאלת תמלוגים כדי לזהות הישגים ולדרבן נציגים כשהישגים אלו מרשימים. בצורה זו ניתן אף לדרבן עובדים להשיג הכשרה חריגה כמו לימוד שפות או השגת תואר אקדמאי כזה לעיתים קיימת מחלוקת חריפה בין נציגי השירות לבין המנהלים הקרובים אליהם בהקשר של יכולת הערכה של המנהלים. קודם הערכות נציגים התבססו על הערכת הביצועים מניתוח מספרים שצינו את כמות הלקוחות שכל נציג שירת ומדדי איכות שלא היו ברורים לנציגים.

שיטת ה SBR מספקת תמונה חדה וברורה של כישורים נחוצים ורכישת מיומנויות לקידום.

ה "מפה" הזו מספקת שני יתרונות עיקריים:

- הקריטריון להערכה ניתן לתיאור והגדרה באופן הרבה יותר מפורט, מה שמאפשר להערכות להיות הרבה יותר חד משמעיות לעובד המצוי תחת ביקורת.
 - הנציג מודע לעבודה הנדרשת כדי להתקדם לשלב הבא ברמת הכישורים שלו ומכאן הוא מסוגל לקשר באופן ברור בין כישוריו לתגמולי השכר והתנאים שהוא מקבל.
- היכולת למנוע טענות הגורסות כי קיימת הפליה בין העובדים או הטיה של הערכות היא חשובה ביותר ליצירת אוירה נעימה בארגון. המקום בו עבודה קשה מתוגמלת כראוי וקידום הקריירה מתואר באופן ברור יש פחות סיכון שדעתו של נציג תוסח בזמן הערכתו או בזמן מתן השירות. לכן ה SBR מתקבלת ע"י העובדים כמערכת התומכת בהם ופועלת לטובתם.

3.13 סיכום -

כל פרויקט SBR מתחיל עם אסטרטגיה המגדירה איך האנשים, תהליכים וטכנולוגיה ישתלבו יחד. שילוב זה יקבע כיצד יש לשנות ולעדכן את האסטרטגיה.

פרויקט ה SBR הוא תהליך מתמשך שבו כל מרכיב משפיע, ומושפע מהמרכיבים האחרים. כאשר התכנון הוא מושכל ומוטמע באופן יעיל ע"פ הכללים העסקיים של הארגון, ה SBR מאפשרת לארגון לדבר עם הלקוחות בקול אחד, מאפשרת לשרתים להתפנות ולטפל באופן יותר פרודוקטיבי בלקוחות המועדפים ומעניקה כוח למנהלים להתמקד במשימות הליבה של ניהול המשאבים, תהליכי למידה ותהליך איסוף המידע.

מערכת ה SBR המושרשת בצורה יעילה, מפיקה תועלת לכל הנוגעים בדבר. לקוחות מקבלים שירות מהיר ויעיל- לא עוד זמני המתנה ארוכים, או טרטור בין שרתים רבים. העסק מצליח ליצור שירות איכותי יותר, מגדיל את נאמנות הלקוחות, ומגדיל את אפשרויות הצמיחה לטווח ארוך.

פרק 4

סקירה ספרותית

4.1 "Telephone Call Centers: Tutorial, Review, and Research

Prospects" By: Noah Gans, Ger Koole, Avishai Mandelbaum (2003)

המאמר הנוכחי מסכם את אופי התפקוד של המוקדים הטלפוניים מבחינת כל האספקטים הנוגעים לניהול נכון של מוקד שירות. לאחר סקירה המאמרים שנכתבו בעבר בנושא, מאמר זה מספק הסבר לגבי מחקרים נוספים שצריכים להיעשות על מנת לשפר את ניהול המוקדים הטלפוניים בעתיד.

מערכות שירות טלפוניות הינן חלק אינטגרלי מעסקים רבים כיום. מוקדי שירות טלפוניים נהפכו לאמצעי תקשורת מועדף של ארגונים עם הלקוחות שלהם. פער הביצוע של המערכות הטלפוניות הינו הולך וגדל וזאת עקב המורכבות ההולכת וגדלה - כגון: פילוח לקוחות רב, מספר שרתים רב, מספר קווים רב וניתוב מבוסס מיומניות.

ישנם שלושה מודלים שבעזרתם ניתן לחזות את כמות הנציגים הדרושים כל פרק זמן. כל אחד מהמודלים מניח כמות לקוחות אינסופית, אקראיות ההגעה וההמתנה בתור.

Erlang C : הוא המודל שהינו הכי נוח לשימוש אך הכי פחות מדויק מבחינת התייחסות לכל המדדים הנדרשים בעת ביצוע החיזוי כמו הנטישות וחסימת הקווים.

Erlang B : מודל זה מוריד את מספר השיחות הממתנות ע"י הוספת קווים, כלומר – הוספת תופעת ה-blocking, אך ללא התחשבות בנטישות.

Erlang C : זהו המודל המוסיף על קודמיו התחשבות בקבלת צליל תפוס ובנטישות.

כמו כן, ישנם שלושה סוגים של אסטרטגיות לניהול מערכות שירות טלפוניות :

1. אסטרטגיה מכוונת יעילות.
2. אסטרטגיה מכוונת איכות .
3. שילוב בין הראשונה לשנייה – QED.

לפי אחוז הממתינים ניתן לדעת את מכוונות המוקד.

- במערכת מכוונת יעילות אין נטישות והסבלנות הינה גבוהה כי בממוצע משך ההמתנה יהיה כחצי מזמן שירות ממוצע, ולכן כמעט כולם ימתינו.
- במערכת מכוונת איכות הרוב יקבלו שירות מייד.
- במערכת מכוונת שירות ויעילות כחצי מהלקוחות יקבלו שירות מייד והחצי השני ימתין בסדר גודל שהינו פחות ממשך השירות. במערכת זו המשאבים ינוצלו ברמה גבוהה- קרוב ל – 90%. היתרון לגודל מאפשר להשיג שירות ויעילות בו זמנית.

הדרך לטיפול בסוגים שונים של התמחויות מבוסס על האסטרטגיות בהן נרצה לנקוט. האחת הינה כי כל הנציגים יתמכו בכל סוגי ההתמחויות השונות. דרך זו לטיפול הינה טובה עבור מוקדים בהם יש מעט התמחויות. הדרך השנייה הינה להכשיר נציגים מסוימים להתמחויות מסוימות. דרך זו לטיפול הינה טובה עבור מוקדים בהם קיימים מספר רב של התמחויות שונות.

ישנם שני משטרים על פיהם מבוסס ניתוב מבוסס התמחויות:

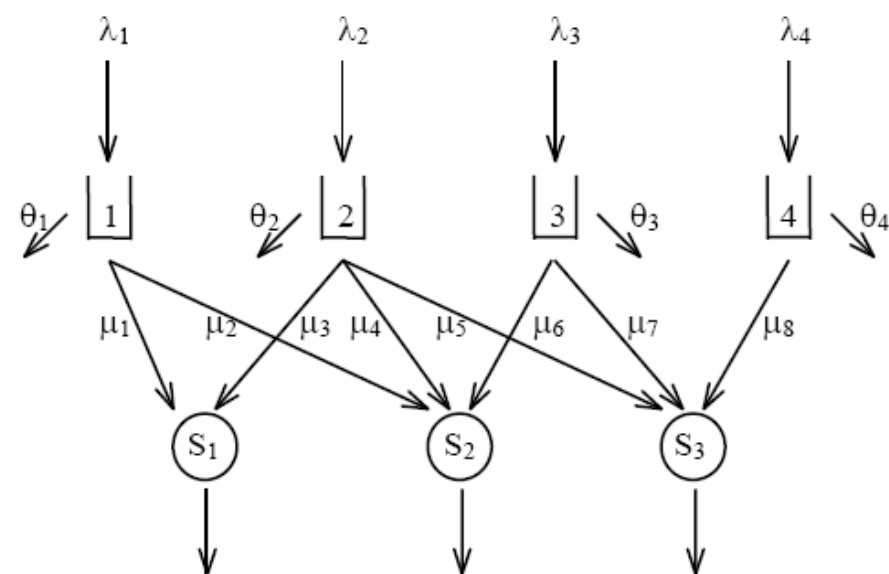
- ניתוב להתמחויות שונות במשטר מבוסס יעילות: הנציגים מגיעים ליעילות של 100%, בדרך שבה בגבול מסוים כל הלקוחות ממתינים בתור.
- ניתוב להתמחויות שונות במשטר מבוסס יעילות ואיכות: בעיית התור הנוצר במשטר הקודם, הופכת במשטר זה לבעיית שליטה בהפצה, כלומר התור הופך לבעיה שקשה לשלוט מי ממתין וכמה זמן.

לסיכום, רוב המודלים התפעוליים שמתחשבים בהגעות, שירות, נטישה - הינם פונקציות של התנהגות אנושית שאותה קשה לחזות ולהגיע למודל אנליטי מובנה. מחקרים נוספים בנושא זה יעזרו לבנות מודל אנליטי יותר מדויק בעתיד.

4.2 "An Introduction to Skills-Based Routing and its Operational Complexities." By Ofer Garnett and Avishai Mandelbaum (2003)

הקדמה:

ניקח בחשבון את המערכת בתרשים 1:



תרשים 1

λ_i - קצב ההגעות לתור i .

μ_i - קצבי השירות

θ_i - קצבי העזיבה מתור i .

S_i – מספר השרתים בכל מאגר.

מערכת זו היא מערכת המבוססת על ניתוב מבוסס מיומנות, כיוון שכל תור מייצג בעצם לקוחות הדורשים סוג מסוים של שירות, ולכל מאגר שרתים מוגדרת מיומנות אותם השרתים מסוגלים לבצע. בתרשים 1, החיצים הנכנסים למאגר מסוים מגדירים בעצם את סוג השירות המבוקש. לדוגמא – שרת ממאגר 2 יכול לשרת לקוחות מסוג 3 בקצב שירות - μ_6 לקוחות ליחידת זמן.

כאשר מיישמים מערכת מבוססת מיומנות יש לבצע מספר החלטות:

1. מי הם הלקוחות?

החלטה זו אינה מתקבלת online והיא בד"כ נעשת ע"י מחלקת השיווק של הארגון. ניתן לחלק את הלקוחות ע"פ סוגי השירות שהם מבקשים (לדוגמא תמיכה בשירותי אינטרנט או בקשה לביצוע פעולות בניירות ערך) או על פי העדיפות שלהם (לדוגמא לקוחות זהב מול לקוחות רגילים).

2. מי הם השרתים?

יש להגדיר את המיומנות של השרתים ואת מספר השרתים בכל קבוצת מיומנות. החלטה זו גם היא אינה מתקבלת online ובד"כ נעשית במחלקת כ"א ו/או במחלקת תפעול. ניתן לחלק את השרתים על פי היכולות השונות שלהם (לדוגמא שרתים דוברי שפות שונות), על פי ההכשרה שלהם (שרת ממוחה לעומת שרת כללי) או על פי הניסיון שלהם (לדוגמא שרת וותיק לעומת שרת טירון).

3. כיצד לנתב את הלקוחות לשרתים?

החלטה זו כן מתקבלת online בעת הגעת הלקוח למערכת, כאשר הפרוטוקול לניתוב נקבע ע"י מערכת התפעול. החלטה זו מוחלטת משתי תתי החלטות:

1. איזה לקוח ינותב לשרת שהתפנה משירות.

2. באם לקוח מגיע למצב בו יש מספר שרתים פנויים, לאיזה שרת לנתב אותו.

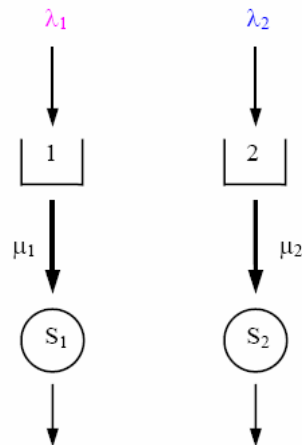
ניתוב מבוסס מיומנות הוא פרוטוקול הניתוב בזמן אמת של הלקוחות השונים לשרתים השונים. החלטת הניתוב יכולה להיות בשני מצבים:

1. סטטאטית – קבועה לאורך הזמן ולא משתנה בהתאם למצב המערכת.

2. דינאמית – החלטה המשתנה עם הזמן בהתאם למצב המערכת (לדוגמא – על פי אורך התור).

מודל N עם שרת אחד בכל מאגר שרתים :

נסתכל על מוקד שירות שבו שני שרתים המספקים תמיכה טכנית עבור שני מוצרים. המודל המתקבל במקרה זה הוא מודל I^2 :



כל שרת מתמחה במוצר יחיד ומשרת לקוחות המבקשים תמיכה טכנית עבור מוצר זה בלבד. נניח כי לקוחות מסוג אחד הם לקוחות הליבה של המערכת. הם נחשבים כלקוחות חשובים יותר מלקוחות אחרים ולכן ישורתו על ידי שרת מומחה – שרת 1.

ניתן לנתח את המערכת כשתי מערכות $M/M/1$, עבור הנתונים הבאים :

$$\lambda_1 = 1.3 \cdot \rho$$

$$\lambda_2 = 0.4 \cdot \rho$$

$$\mu_1 = \mu_2 = 1$$

$$S_1 = S_2 = 1$$

כאשר ρ הוא עומס התנועה במערכת (מספר דקות השירות הנכנסות למערכת לדקה).

כאשר ρ קטן אזי המערכת לא עמוסה ולכן ביצועי המערכת כולה יהיו משביעי רצון, כפי שניתן לראות בטבלה 1:

Table 1 : Performance of I ² design						
Traffic Load (p)	Type 1 Customers / Server 1			Type 2 Customers / Server 2		
	Average Wait (seconds)	Average Queue Length	Utilization	Average Wait (seconds)	Average Queue Length	Utilization
1/3	46	0.3	43%	9	0.0	13%
2/3	390	5.6	87%	22	0.1	27%
3/4	1170	38.0	97.5%	26	0.1	30%

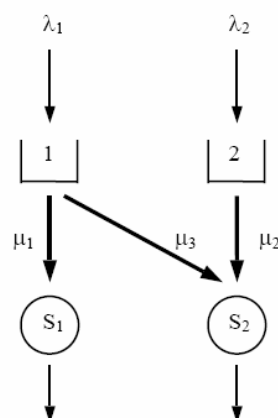
טבלה 1

אולם ניתן לראות מטבלה 1 כי ככל שהעומס גדל השרת הראשון עמוס מאוד (נצילות של 97.5%), ולכן נוצר מצב כי הלקוחות מסוג 1 נאלצים להמתין זמנים ארוכים בתור, בעוד שהעומס על השרת השני אינו גדול כלל (30%).

דבר זה פוגע ביעילות המערכת במיוחד תחת ההנחה שהזכרנו כי הלקוחות מסוג 1 הם לקוחות בלעי חשיבות רבה למערכת.

השאלה הנשאלת היא, האם ניתן לתכנן מערכת חדשה שתהיה יעילה יותר ותבטיח שירות טוב יותר ללקוחותיה. האלטרנטיבה הראשונה לפתרון היא ליישם את מודל V עם מאגר של שני שרתים. צורה זו היא היעילה ביותר, אך חסרונה הוא בכך שהשרת הראשון המתמחה ישירות לקוחות מסוג 2.

האלטרנטיבה השנייה היא מודל N המתואר בתרשים הבא:



מודל זה מנצל את התוצאה ששרת 2 אינו עסוק כמו שרת 1 ולכן הוא יוכל לעזור לשרת 1 ע"י מתן שירות ללקוחות מסוג 1. מודל זה מצריך הכשרה נוספת של שרת 2 על מנת שיוכל לשרת לקוחות מסוג 1. הנחה הגיונית במקרה זה היא שהשירות שניתן ללקוחות 1 ע"י שרת 2 יהיה איטי יותר, ולכן במקרה זה הנתונים הם :

$$\lambda_1 = 1.3 \cdot \rho$$

$$\lambda_2 = 0.4 \cdot \rho$$

$$\mu_1 = \mu_2 = 1$$

$$\mu_3 = 0.5$$

$$S_1 = S_2 = 1$$

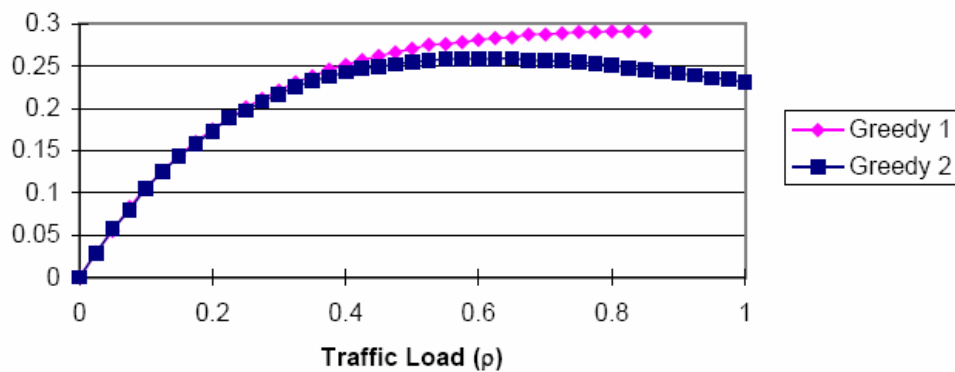
נבחין בין 2 תסריטים :

- 1.1 greedy – שרת 2 ישרת לקוחות מסוג 1 כל עוד ישנם בתור לקוחות מסוג 1 ושרת 1 עסוק.
2. greedy 2 – שרת 2 ישרת לקוחות מסוג 2 כל עוד ישנם לקוחות מסוג 2 המחכים בתור, אחרת ישרת לקוחות מסוג 1 המחכים בתור, אם שרת 1 עסוק.

בשני התסריטים המוצגים לעיל כאשר השרתים כולם בטלים לקוחות מסוג i ינותבו לשרת מסוג i .

מהי כמות העזרה ששרת 2 נותן לשרת 1?

Graph 1 : Fraction of Type 1 served by Server 2



- כאשר העומס הוא נמוך ישנה מעט עזרה לשרת 1 משום ששרת 1 יכול בקלות להתמודד עם לקוחות מסוג 1.

- כאשר העומס שווה ל-1/3 אנו מתחילים לראות את השינוי בין כמות העזרה שנותן השתר השני לראשון בשני התסריטים. התוצאה היא שתחת greedy 1 השרת הראשון מקבל הרבה יותר עזרה.
- ניתן לראות כי ככל שהעומס עולה כמות העזרה תחת greedy 1 שמקבל שרת 1 גדלה.
- תחת greedy 2 ניתן לראות כי כשהעומס עובר את ה-80% לערך כמות העזרה ששרת 2 מעניק לשרת 1 קטנה שכן שרת 2 מטפל בלקוחות 2.

ניתוח יציבות המערכת:

ניתן לראות כי שני הגרפים המוצגים בגרף 1 אינם מסתיימים עבור אותו הערך של העומס. מסתבר שמערכת הפועלת על פי greedy 1 אינה יציבה לכל $\rho < 1$ כפי שהיינו מצפים.

נבחן את המקרה בו $\rho = 1$:

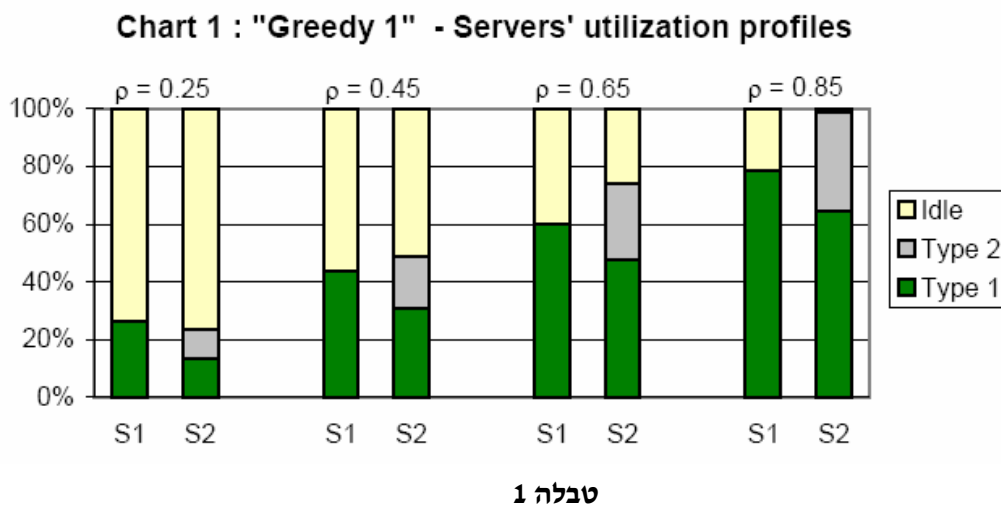
אם שרת 1 תורם את כל זמנו על מנת לשרת לקוחות מסוג 1, שרת 2 יצטרך לשרת: $\lambda_1 - \mu_1 = 1.3 - 1 = 0.3$ לקוחות לדקה. לקוחות אלו דורשים $\frac{0.3}{\mu_3} = \frac{0.3}{0.5} = 60\%$ מזמנו של

שרת 2. 40% הזמן הנותרים מספיקים בדיוק על מנת לשרת את הלקוחות מסוג 2 כיוון ש- $\frac{0.4}{\mu_2} = \frac{0.4}{1} = 40\%$. אולם, כפי שראינו בגרף 1 המערכת תחת greedy 1 מתפוצצת כבר עבור

עומס ששווה לכ-90% ואילו המערכת תחת greedy 2 יציבה תמיד.

האם אפשר להסביר מדוע המערכת ב-greedy 1 מתפוצצת לפני $\rho = 1$?

נסתכל על טבלה 1:



אם נסתכל על המצב בו העומס שווה ל-85% נראה כי שרת 2 כמעט תמיד עסוק בעוד ששרת 1 פנוי ביותר מ-20% מזמנו. יתרה על כך, שרת 2 מבילה את רוב זמנו בטיפול בלקוחות 1. המסקנה הנובעת מתרשים זה היא ששרת 2 עוזר לשרת 1 יתר על המידה ולכן מזניח את לקוחות מסוג 2 עד למצב בוא אורך התור שואף לאינסוף. למעשה שרת 2 עובד קשה מדי ע"י כך שהוא לקוח משרת 1 לקוחות ששרת 1 היה צריך להתמודד איתם לבדו, וזהו הגורם לפיצוץ המוקדם של המערכת. בשלב זה לא ניתן לגעת מראש עבור איזה ערך מדויק של העומס המערכת תתפוצץ.

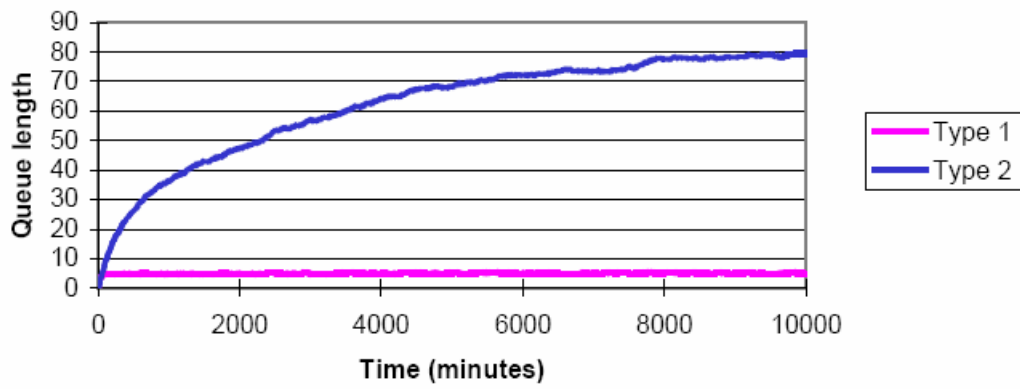
ראינו כי **greedy 1** מכתוב חלוקה לא יעילה של העומס במערכת, ועל ידי כך מונע מהמערכת להגיע לפוטנציאל עבודה מקסימאלי. אף על פי כן מערכת זו אכן מספקת את הדרוש מבחינת רמת השירות הנדרש עבור לקוחות מסוג 1 שהוגדרו כלקוחות הליבה של המערכת, מכיוון שהיא נותנת עדיפות רבה ללקוחות אלו. ב-**greedy 2** לא ניתנת עדיפות זו, ולכן נתבסס מעתה והלאה על **greedy 1** בלבד.

על מנת להתגבר על הבעיה הנוצרת כתוצאה ממדיניות **greedy 1** נציג מדיניות בקרה באמצעות סף. זוהי למעשה וריאציה שונה של **greedy 1** עם ההגבלה הבאה: שרת 2 יעזור לשרת 1 רק אם מספר האנשים בתור 1 גדול מסף מסויים.

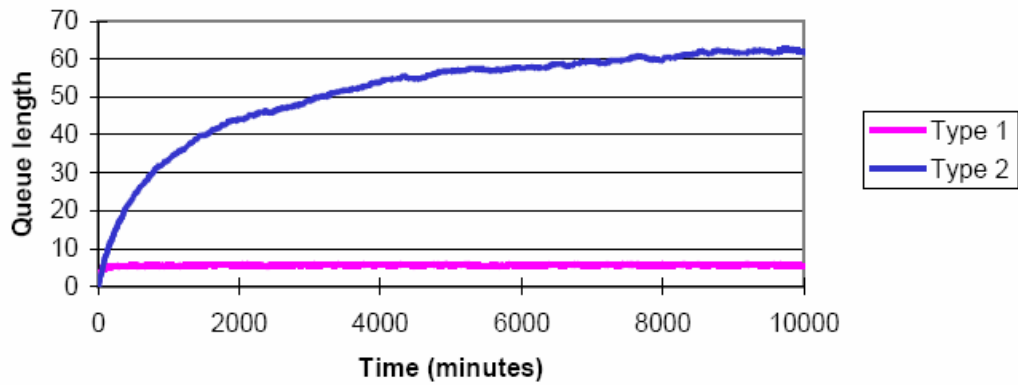
חשוב לראות כי מדיניות זו דורשת להגדיר מראש מהוא הסף. כאשר הסף הוא 1 מדובר למעשה במדיניות **greedy 1** המקורית. מדיניות זו אינה משמרת עבודה (**work conserving**), כלומר, יכול להיות מצב שלקוחות מסוג 2 יחכו בתור וזאת למרות שקיימים שרתים פנויים מסוג 2. לדוגמא במקרה שהסף נקבע על עשרה שרתים פנויים וקיימים רק 4 שרתים פנויים, הם לא ישרתו את הלקוחות מסוג 2 הממתינים בתור עקב מדיניות הסף.

עבור עומס של $\rho = 0.95$ יישום מדיניות הסף יכולה להביא למערכת יציבה.

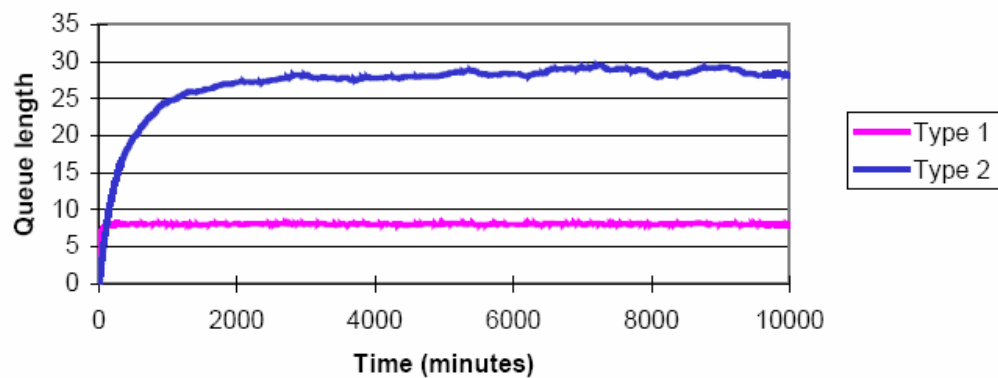
Graph 4 : Threshold = 4 , $\rho = 0.95$ - Queue Lengths



Graph 5 : Threshold = 6 , $\rho = 0.95$ - Queue Lengths



Graph 6 : Threshold = 8 , $\rho = 0.95$ - Queue Lengths



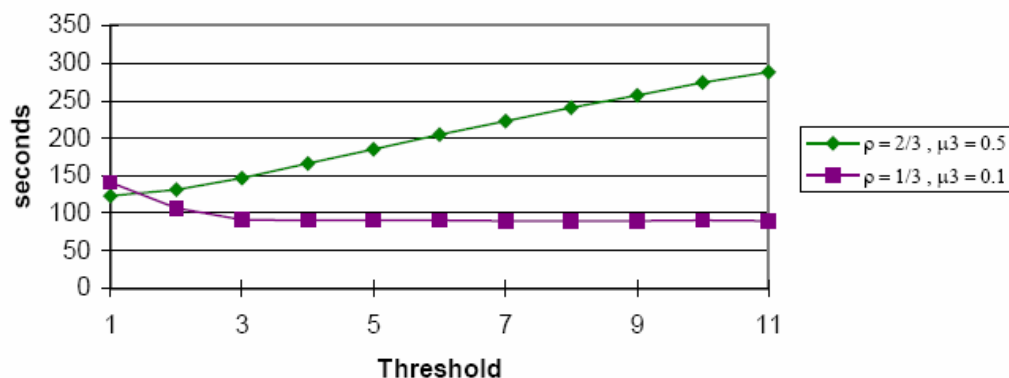
בגרפים 4-6 אנו רואים כי אורך התור של לקוחות מסוג 2 לא ישאף לאינסוף ולכן המערכת במצב יציב. בנוסף, ע"י השוואה בין הגרפים ניתן לראות כי כאשר הסף גבוה יותר אזי התור של לקוחות מסוג 1 גדל ואילו התור של לקוחות מסוג 2 קטן.

ע"י מדיניות הסף הנקובה לעיל אנו מפחיתים את כמות העזרה של שרת 2 לשרת 1, ועבור סף גבוה דיו לשרת 2 יכול להיות מספיק זמן על מנת לשרת את לקוחות מסוג 2, וע"י כך למנוע את התפוצצות המערכת.

חשוב לציין כי הסף יכול להיות מבוסס גם כן על תור הלקוחות מסוג 2. במקרה זה שרת 2 יעזור לשרת 1 רק כאשר מספר הלקוחות מסוג 2 הממתינים בתור הוא מתחת לסף כלשהו. הבעיה עם שימוש במדיניות זו היא שהיא אינה אפקטיבית כמו קודמתה. הרעיון שעומד מאחורי שלילת מדיניות זו הוא שיש להתייחס בסף לתור המשותף לשני השרתים. בנוסף למדיניות זו חסר המאפיין הקריטי של מניעה משרת 2 לעזור לשרת 1 כאשר העומס על שרת 1 קטן.

זמן השהייה במערכת:

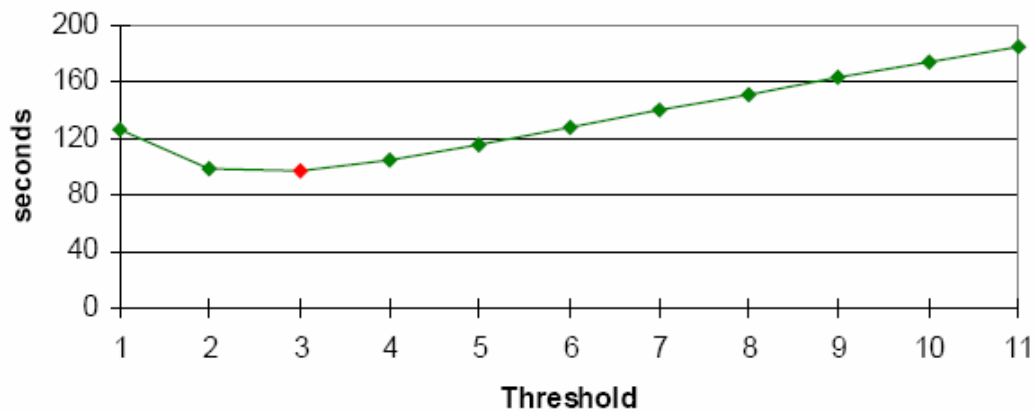
Graph 8 : Type 1 - Average Time in System



התוצאות המובאות בגרף 8 מראות כי עבור $\rho = \frac{2}{3}$ ככל שהסף גבוה יותר, לקוחות מסוג 1 ישהו במערכת זמן רב יותר. אולם הסקה זו לא נכונה עבור כל מקרה שכן עבור סט הפרמטרים השני המוצג בגרף נקבל את אפקט "השרת האיטי" מכיוון ששרת 2 יהיה איטי פי 10 משרת 1, ולכן לקוחות מסוג 1 שישורתו ע"י שרת 2 ישהו זמן רב במערכת, ולכן ככל שהסף עולה – מספר הלקוחות מסוג 1 שישורתו ע"י שרת 2 ירד ולכן זמן השהייה הממוצע יתקצר. בעצם בדוגמא זו הנזק שנגרם ע"י השרת האיטי הוא גדול יותר משה"כ התועלת של הקטנת זמני ההמתנה עבור לקוחות מסוג 1.

מגרף 9 ניתן לראות כי קיים סף אופטימלי עבורו סה"כ זמני ההמתנה בתור הוא מינימלי:

Graph 9 : $\rho = 2/3$ - Overall Average Wait in Queue



חשוב לזכור כי את העדיפות של לקוחות מסויימים על לקוחות אחרים ניתן לתרגם למונחים כספיים. לדוגמא, במקרה בו עלות ההמתנה של לקוח היא פרופורציונאלית לזמן ההמתנה שלו בתור. כמו כן נזכור כי עבור לקוחות מסוג 1, שלהם חשיבות רבה יותר, עלויות זמן ההמתנה שלהם גדולה מאשר עלויות ההמתנה של לקוחות מסוג 2.

בהתאם לכך נוכל למצוא את הסף האופטימלי המביא למינימום את סה"כ העלויות. כמובן שסף זה לא יביא בהכרח למינימום זמן המתנה עבור שני סוגי הלקוחות.

כיצד משתנה הסף האופטימלי כתלות בקצב ההגעה של לקוחות 2?

ככל שקצב ההגעה של לקוחות מסוג 2 גדל כך גדל הסף האופטימלי המביא למינימום עלויות. דבר זה עקבי עם הפרוש לסף האופטימלי כנקודת איזון בין השארת שרת 1 ללא עבודה לבין הזנחת הלקוחות מסוג 2, ולכן, ככל שקצב ההגעה של הלקוחות גדל נצטרך סף גבוה יותר ע"מ לא לזנוח את הלקוחות מסוג 2.

חשוב לראות כי ניתן לקבוע את מדיניות הסף כאשר הסף הוא לא על מספר הלקוחות הממתינים בתור אלא על מספר השרתים הפנויים מסוג 2.

4.3 "RESOURCE POOLING AND STAFFING IN CALL CENTERS WITH SKILL-BASED ROUTING". By: Rodney B. Wallace and Ward Whitt(2004)

1 - הקדמה:

מטרת המאמר היא לספק שיטה האמורה לשפר את התיכון והניהול של מוקדי שירות טלפוני או באופן כללי יותר מרכז קשרי לקוחות המאפשרים התקשרות דרך מדיות אחרות כמו דואר אלקטרוני. תחילה נסקור מספר מושגים עיקריים:

ניתוב מבוסס מיומנות -

מוקדי שירות בד"כ מטפלים במספר סוגים של לקוחות. עם זאת בד"כ זה לא אפשרי או יעיל מבחינת עלות, להכשיר כל נציג שירות כך שיוכל לטפל בכל סוגי הלקוחות. לדוגמא:

1. מוקד רב לשוני – הן השרתים והן הלקוחות דוברים בשפות שונות, וכמובן האפשרות של להכשיר את כל השרתים לדיבור בכל השפות אינו אפשרי.

2. קיימים מוקדים המטפלים בשיחות אשר משלבות מספר יוזמות מיוחדות.

3. מוקדי תמיכה טכנית.

למעשה, לשיחות שונות יש דרישות שונות ולשרתים שונים יש מיומנויות שונות. מנתבי שיחות אוטומטים – (ACD) מסוגלים כיום לכוון שיחות לנציגים בהתאם למיומנויות הנדרשות. הציוד של טכנולוגיות המידע והתקשורת של מוקדי השירות יכולים לאפשר הכללה של שיחות מסוגים שונים. היכולת הזאת נקראת ניתוב מבוסס מיומנות (SBR).

Resource Pooling – איגוד משאבים

איגוד משאבים הוא פתרון בין שני מילים לבעיית ה-SBR.

במשמעות של SBR, איגוד משאבים הוא שעם הגבלת מספר המעברים האפשריים ואלגוריתם שיבוץ הגיוני, מוקד שירות המבוסס על SBR יכול לעבוד ברמה הקרובה מאוד לזו של מוקד בו לכל השרתים יש את כל הקישורים הנדרשים.

איגוד משאבים הוא חשוב במקרה זה שכן הוא מוביל ליעילות גבוהה הרבה יותר, עם ביצועים ותועלות גבוהות יותר, והורדה בעלויות. כמובן שאיגוד משאבים הוא אינו הפתרון לכל הבעיות, אבל כוונת המאמר היא להראות שזהו רעיון חשוב עבור מוקדי שירות מבוססי SBR.

1. ללא ספק המשמעות העיקרית של תוצאות מחקר זה הוא שיש לשים לב בזהירות לשילוב הכישורים של השרתים השונים.
2. ניסוי זה מראה כי יכולה להיות תועלת רבה בגמישות מסוימת (שרתים בעלי כישורים כפולים), אך הערך המוסף של גמישות רבה (שרתים בעלי מספר גדול של כישורים) הוא לא בהכרח גדול.
3. לאיגוד משאבים יש משמעות רבה על האלגוריתמים של SBR.
4. אם אכן גמישות מוגבלת בשילוב עם מסגרת עבודה סאב-אופטימאלית של SBR יכולה להביא לביצועים טובים כמעט כמו במקרה של שרתים יוניברסליים, אז נצליח לעקוף בהצלחה בעיית שיבוץ קשה במיוחד.
5. אם נתכן וננהל מוקד שירות מבוסס SBR כך שתהיה בו גמישות מספקת על מנת לתמוך באיגוד משאבים, יהיה קל יותר לתפעל את מוקד השירות. במאמר זה אנו מראים איך לנצל איגוד משאבים במטרה לפתח אלגוריתם שיקבע את האיוש והציוד (מספר קווי הטלפון) הדרושים במוקד שירות מבוסס SBR. על מנת לעשות זאת נציג מסגרת עבודה ספציפית המיישמת SBR. וכך במסגרת העבודה נפתח אלגוריתם לאיוש והספקת הציוד.

A little Flexibility Goes a Long Way

הנקודה המרכזית היא: *A little flexibility mat go a long way*. המאמר יראה כי מוקד שירות מבוסס SBR בו לכל שרת מספר מצומצם של כישורים, ובפרט שני כישורים, יכול לבצע כמעט טוב כמו מוקד שירות מבוסס SBR בו כל השרתים הם יוניברסליים. המחקרים התיאורטיים המשמעותיים שגילו את היתרון הגדול של מעט גמישות היו של Azar (1994) ו-Vvedenskaya (1996), Turner (1996) ו-Mitzenmacher (1996). על מנת להעביר את הרעיון, המאמר מתאר את בעיית התורים המתוארת אצל Vvedenskaya (1996). במודל המתואר בבעיה יש זרם אחיד של לקוחות המגיעים בתהליך פואסון. עם ההגעה כל לקוח חייב להצטרף לאחד ממספר רב של תורים בלתי תלויים, זהים, עם שרת יחיד המשרת בקצב אקס' כאשר אורך התור האפשרי הוא אינסופי. הגישה הסטנדרטית של בעיית ההצטרפות לתורים היא בדיקת המצב של כל תור והצטרפות לתור בו יש הכי פחות לקוחות. אכן Winston (1977) מצא כי הגישה הזו היא אופטימאלית, אך הקושי עם גישה זו הוא שהיא דורשת כמות רבה של מידע. דרך אחת להפחית את כמות המידע הנדרשת, היא להשים כל לקוח שמגיע רנדומאלית באחד מהתורים, כאשר לכל תור יש הסתברות שווה להצטרפות. בד"כ זו אין צורך במידע כלשהו, אך השמה רנדומאלית שכזו גורמת לכל שרת לפעול כ- $M/M/1$ ואז אנו מאבדים את היתרון לגודל.

רעיון המפתח של Azar (1994), Vvedenskaya (1996), Turner (1996) ו-Mitzenmacher (1996) היא להקטין את תחום ההחלטה. אלטרנטיבה פשוטה יותר היא לבחור שני תורים רנדומאלית, ולהצטרף לתור הקצר מבניהם. העובדה המדהימה בכך היא שמצד אחד אין צורך בכמות המידע רב, ומהצד השני האלגוריתם מביא לביצועים דומים כמו במקרה של הצטרפות לתור הקצר ביותר, הדורשת מידע מלא. המסקנה נעשתה רשמית ע"י שימוש בהתנהגות גבולית על מספר השרתים (ושמירה על קצב הגעה פרופורציונאלי).

2 – מודל מוקד שירות טלפוני.

סימונים:

- C – מספר השרתים.
- K – קווי המתנה נוספים. (סה"כ $K+C$ קווי טלפון)
- n – מספר סוגי הלקוחות.

הנחות:

- קו טלפון יכול להיות מנוצל ע"י מתן שירות או ע"י לקוח ממתין.
- ה-ACD יכולה להכיל $C+K$ שיחות אם אין קווים המנוצלים להמתנה.
- שיחות המגיעות כאשר כל $C+K$ הקווים תפוסים נחסמות ואובדות.

הגדרת מסגרת העבודה ל-SBR :

- לכל שרת בין 1 ל- n כישורים.
- שרת יכול לתת שירות לשיחה מסוג מסוים אם הוא בעל הכשרה מתאימה לסוג שירות זה.

2.1 השמת כישורים לשרת –

נגדיר מטריצת כישורים לשרת – A - בגודל $C \times n$. כל שורה מתאימה לשרת ואילו כל עמודה מתאימה לרמת מיומנות (משתמשים ברמות מיומנות על מנת לתת עדיפות בין שרתים שונים בעלי אותה מיומנות). לשרת יש מיומנות מסוימת אמ"מ יש לו ערך חיובי במטריצה תחת אותה רמת מיומנות. לרמות מיומנות נמוכות יש העדפה בסכמת העדיפויות.

הטור הראשון ב- A מציין את המיומנות הראשית. ההנחה היא שלכל שרת יש מיומנות ראשית.

$$\text{כלומר: } 1 \leq A_{i,1} \leq n \quad \forall i$$

נקבץ את השרתים לקבוצות עבודה על פי המיומנות הראשית שלהם, כך שקבוצה G_k היא תת קבוצה

$$\text{של השרתים עם מיומנות ראשית } k - G_k = \{i : 1 \leq i \leq C, A_{i,1} = k\}, \quad 1 \leq k \leq n$$

הנחות:

- מיומנות של שרת מופיעה לכל היותר ברמה אחת.
- לכל שרת יש מיומנות אחת בלבד בכל רמת מיומנות.
- שני שרתים הם זהים (יש להם את אותן מיומנויות באותן רמות מיומנות) אם"מ השורות שלהם במטריצה A זהות.

2.2 מה לעשות בעת הופעת לקוח?

- כאשר מגיעה שיחה מסוג k היא מנווטת תחילה לשרת פנוי מקבוצת עבודה k , שכן ההתמחות הראשית של שרתים מסוג זה הוא סוג שירות k .
- על מנת להחליט לאיזה מהשרתים הפנויים בקבוצה k נפנה את הלקוח שהגיע, נשתמש במדיניות – $LIAR$ - longest idle agent routing. ע"פ השיטה הלקוח יופנה לשרת מקבוצה k שנמצא הכי הרבה זמן במצב מנוחה.
- אם כל השרתים בקבוצת עבודה k עסוקים, אזי השיחה מועברת לשרת פנוי בעל מיומנות k כמיומנות משנית. שוב מבין כל השרתים הפנויים בעלי מיומנות משנית k נבחר את השרת ע"י שיטת $LIAR$. אם כל השרתים בעלי מיומנות משניים k תפוסים נעבור לשרתים בעלי מיומנות k ברמה שלוש וכן הלאה.
- אם אף שרת בעל מיומנות K בכל רמה אינו פנוי, לקוח k יועבר לסוף התור המשוך לקבוצת עבודה k .

2.3 מה לעשות כאשר מתפנה שרת?

- אם אין לקוחות ממתינים באף אחד מ- n התורים השרת הופך להיות פנוי.
- אחרת – השרת בודק בכל התורים של קבוצות העבודה עבורן הוא מיומן. אם אין לקוחות באף אחד מהתורים הוא הופך להיות פנוי. כלומר – השרת בודק את התורים לפי רמת המיומנות שלו (כלומר תחילה הוא הולך לתור של המיומנות הראשית שלו, לאחר מכן למשנית וכן הלאה). השרת משרת את הלקוח הראשון בתור, הנמצא בתור הראשון שאינו ריק – כלומר – השרת משרת פניות בכל קבוצת עבודה לפי מדיניות $FCFS$, כל עוד הוא יכול לשרת אותו.

2.4 הנחות סטוכסטיות –

המאמר מציין מספר הנחות סטוכסטיות הקשורות לסימולציה. ההנחות הן הנחות פשוטות למדי, אבל החוקרים נוטים להאמין כי הן אינן קריטיות לתוצאות:

1. במערכת יש n סוגי שיחות נכנסות. כאשר כל ההגעות מהוות תהליכי פואסון בלתי תלויים עם קצבי

הגעה λ_i , $1 \leq i \leq n$.

2. ההסתברות ששיחה שתגיע היא $p_i = \lambda_i / \lambda$.

3. זמן השירות הוא מ"מ אקס' ב"ת בהגעות. ממוצע זמן השירות של שיחה מסוג k הוא $1/\mu_k$.

4. במקרה זה אין התייחסות לשיחות חוזרות.

5. מניחים כי רמת השירות גבוהה דיה על מנת להניח כי ניתן להתעלם מנטישות

3 – איגוד משאבים – ניסוי.

בחלק זה המאמר מתאר סימולציה שנערכה החוקרת את גודל ההשפעה שיש לאיגוד משאבים במוקד שירות טלפוני. למעשה חוקרים בכמה שרתים מיומנים יש צורך על מנת ליהנות מהיתרונות של איגוד משאבים.

- הסימולציה מבוססת על מוקד שירות טלפוני בו יש 6 סוגי שיחות נכנסות, ובודקת מה קורה כאשר לכל השרתים יש j כישורים, כאשר j נע בין 1 ל-6 (עבור כל j הורצה סימולציה נפרדת).
- כאשר $j=6$ כל שרת הוא אוניברסלי והמוקד מתנהג כמו מוקד בעל קבוצה אחת.
- כאשר $j=1$ המוקד לא מתנהג בדיוק כמו 6 מוקדים נפרדים, וזאת בגלל מרחב ההמתנה האינסופי (כל השיחות הנכנסות ממתינות בתור אחד ולא בשישה נפרדים). התור היחיד מביא לביצועים טובים בהרבה מאשר שישה תורים נפרדים.

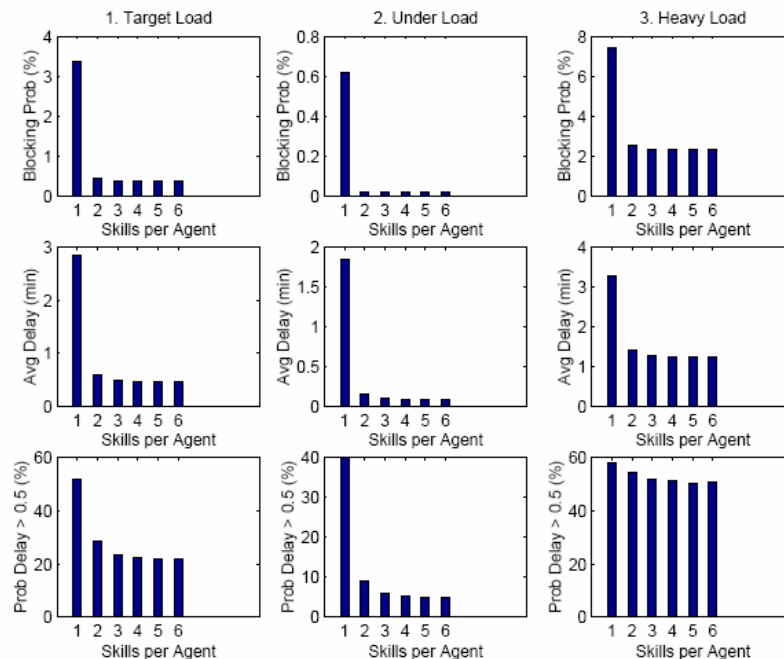
הרעיון הראשי הוא שהביצועים עבור $2 \leq j \leq 5$ קרובים מאוד לאלו המתקבלים עבור $j=6$, והרבה יותר טובים מאשר אלו המתקבלים עבור $j=1$. מכאן ניתן לראות כי איגוד משאבים משמעותי הוא גם כאשר לכל שרת שתי מיומנויות בלבד. הביצועים אכן טובים יותר כאשר $j=3$ מאשר כאשר $j=2$, אך רוב היתרונות של איגוד משאבים הוא כאשר $j=2$.

3.1 מאפייני הסימולציה:

- המודל אותו ניקח בחשבון בסימולציה הוא: $M_6/M/90/30$ (ישנם 90 שרתים ועוד 30 קווים נוספים להמתנה).

- ישנם 6 סוגי שיחות נכנסות ומכאן 6 קבוצות עבודה.
- מספר השרתים בכל קבוצה הוא $90/6=15$ ($C_1=C_2=...=C_6=15$).
- זמני השירות הם מ"מ ב"ת ש"ה אקספוננציאלים כך ש- $\frac{1}{\mu_1} = \dots = \frac{1}{\mu_6} = 10 \text{ min}$.
- מספר המיומנויות של כל שרת הוא בין 1 ל-6 בסדר רץ.
- מכיוון שיש 15 שרתים בכל קבוצת עבודה ו- $j \geq 2$ יהיו בדיוק $15/5=3$ שרתים עם כל קומבינציה של מיומנות ראשונית ומשנית.
- נתחשב בשלושה עומסים: 84 – עומס ממוצע, 77.4 – עומס קל ו-90 – עומס כבד.
- עוצמות ההגעה המתאימות הן: $84/90=0.933$ (עומס ממוצע), $77.4/90=0.86$ (עומס קל) ו- $90/90=1$ (עומס כבד).
- מכיוון שאורך התור האפשרי הוא אינסופי ניתן לאפשר עומס גדול מ-1 בדומה למצב של נטיות.
- כל סימולציה מכילה בממוצע 800,000 הגעות לאחר תקופת חימום המביאה את המערכת למצב יציב. תקופת החימום נבחרה כך שתתאים ל-2000 זמני שירות ממוצעים, כלומר בין 20-24% בכל סימולציה.
- על מנת לחשב את רווחי הסמך השתמשו בשיטה של ממוצעי אצוות. כל ריצה חולקה לעשרים אצוות.

3.2 תוצאות הסימולציה:



התוצאות מראות באופן דרמטי את איגוד המשאבים. בשלושת העומסים הנתונים יש שיפור משמעותי כאשר לשרתים יש שתי מיומנויות. בנוסף לזה, ניתן לראות שמרבית היתרונות הם בהוספת המיומנות השנייה. רק שיפורים קלים מושגים ע"י הוספת מיומנות שלישית והלאה.

4 – הוירסיטיקה לאלגוריתם SBR.

בחלק זה ניישם את מאפייני איגוד המשאבים על מנת לפתח אלגוריתם שיחולל דרישות לקביעת רמת האיוש ומספר הקווים במוקד שירות טלפוני מבוסס SBR. הכוונה היא שעלינו לקבוע את הפרמטרים C ו- K במודל $Mn/M/C/K/SBR$ ובמטריצה A , בהינתן פרמטרים ממודלים אחרים ואילוצים שונים על המטריצה A והביצועים של המודל. הרעיון העיקרי הוא שיש להניח שניתן ליישם איגוד משאבים, ואם זאת לבדוק באמצעות סימולציה שהביצועים שמתקבלים הם אכן משביעי רצון ולקבוע החלטות שונות אם לא כך הדבר.

4.1 אילוצים -

המטרה העיקרית היא למזער את C – מספר השרתים הנדרשים, והמטרה המשנית היא בהינתן C למזער את $C+K$ – סה"כ מספר הקווים הנדרש. המטרה היא להביא למינימום את פרמטרים אלו תוך התחשבות בביצועים משביעי רצון, ואילוצים על A – לכל שרת 2 מיומנויות לכל היותר.

נקבע שני מדדי ביצועים :

$$1. \text{ מהירות המענה כרמת שירות} - P(\text{delay}_i < \tau_i | \text{entry}) \geq \delta_i \quad (4.1)$$

כאשר Delay_i הוא זמן ההמתנה במצב יציב של סוג לקוחות i עם פרמטר τ_i ו- δ_i . במקרה

$$\text{שלנו נבחר: } \tau_i = \tau = 0.5, \quad \delta_i = \delta = 0.8 \quad (4.2)$$

כלומר נדרוש כי 80% מכל השיחות יענו תוך ½ דקה.

$$2. \text{ הסתברות לחסימה} - \text{עבור כל סוג לקוחות } i \text{ נדרוש: } P(Q^{(i)} = C + K) \leq \varepsilon_i \quad (4.3)$$

כאשר $Q^{(i)}$ הוא מספר השיחות במצב יציב שלקוח מסוג i פוגש. תחת ההנחה של הגעות

פואסוניות ועל פי PASTA המשתנה המקרי $Q^{(i)}$ מתפלג כמו Q , סה"כ השיחות במערכת

$$\text{בזמן מקרי. במקרה שלנו נקבע: } \varepsilon_i = \varepsilon = 0.005, \quad \forall i \quad (4.4)$$

כלומר נרשה 0.5% חסימות.

בדרך כלל שרתים יקרים יותר מקווי טלפון, כך שאחוז החסימה צריך להיות קטן משמעותית. אילוץ זה בא למנוע מספר גדול מדי של קווי טלפון. הגישה היא לנצל את מאפייני איגוד המשאבים כדי להשיג פתרון התחלתי. לאחר מכן נשתמש בסימולציה על מנת לקבוע התאמות מקומיות והגיע לתוצאה טובה.

4.2 מציאת הזוג ההתחלתי (C, K) –

נשתמש במאפייני איגוד המשאבים ונתחיל במקרה בו לכל השרתים יש את כל המיומנויות הקיימות, כלומר, כאילו יש לנו סוג שירות אחד. כך נוכל לבחור ערכים ל- C ול- K שיעמדו בביצועים הנדרשים ע"פ מערכת $M/M/C/K$. דרכים לעשות זאת פותחו ע"י Massey and Eallace (2004), עבור הפרמטרים הבאים :

$$(4.5) \quad \lambda = \sum_{i=1}^N \lambda_i$$

$$(4.6) \quad \frac{1}{\mu} = \frac{1}{\lambda} \sum_{i=1}^n \frac{\lambda_i}{\mu_i}$$

$$\alpha = \lambda / \mu$$

ידוע ממחקרים קודמים כי במערכת סטוכסטית רבת שרתים, בעומס גבוה מתקיים :

$$(4.7) \quad C \approx \alpha + c_1 \sqrt{\alpha}, \quad K \approx c_2 \sqrt{\alpha}$$

עבור קבועים חיוביים c_1 ו- c_2 מסדר ראשון. אנו יכולים לחזות את הערכים הקריטיים שיהיו.

מכיוון שיש אלגוריתמים מהירים מאוד לפתרון מערכת $M/M/C/K$, קל לבצע חיפוש לערכים C ו- K האופטימאליים. השיטה הטובה ביותר היא לקבוע בהתחלה את $K = \infty$. עבור $K = \infty$ זמן המענה יהיה מונוטוני עם C , כך שקל יהיה למצוא את C שמתאים לאילוף על זמן המענה. לאחר מכן נקבע ערך התחלתי עבור C ואז אחוז הנחסמים יהיה מונוטוני עם K , וכך קל למצוא את K שמתאים לאילוף על אחוז הנחסמים. מכיוון שאנו מקטינים את K אנו מעלים את אחוז הנחסמים וע"י כך מורידים את זמן ההמתנה של הלקוחות, לכן יכול להיות שבמהלך מציאת K נוכל להקטין גם את C . האלגוריתם מבוסס על אתחול זוג C, K והורדה של C בקפיצות של יחידה כך שנמצא את ה- K המתאים המקרב את אחוז הנחסמים בכל שלב לפונקציית המטרה עד שלא ניתן להוריד יותר את C בהתאם לאילוף על זמן ההמתנה. מציאת K המתאים ל- $(C-1)$ הוא קל יחסי כיוון שאנו יכולים להתחיל עם $K+1$ כאשר K מקושר ל- C כלשהו. בדרך זו נחשב את הזוג (C, K) האופטימאלי למערכת $M/M/C/K$. (כלומר להביא את שניהם למינימום תוך עמידה באילוצים).

4.3 קביעת גודל קבוצות העבודה:

לאחר שמצאנו את (C, K) ההתחלתי יש לקבוע את המיומנויות ורמת המיומנויות של C השרתים. תחילה נקבע גודל התחלתי לקבוצות העבודה – כלומר קביעת C_i – מספר השרתים בקבוצת עבודה i לכל $i, 1 \leq i \leq n$. שוב נשתמש בנוסחאות גבוליות של תהליכים סטוכסטים בעומס גבוה ובקירוב השורש לקביעת מספר השרתים בקבוצות העבודה. באופן פרטני, נתעלם תחילה מאילוצים אינטגרלים ונחולל מספר R_i – הערכה ראשונית ל- C_i .

$$(4.8) \quad R_i = \alpha_i + x\sqrt{\alpha_i}, \quad 1 \leq i \leq n$$

$$(4.9) \quad \alpha_i = \lambda_i / \mu_i$$

$$(4.10) \quad x = \frac{(C - \alpha)}{\sum_{i=1}^n \sqrt{\alpha_i}} \quad \text{כאשר } x \text{ הוא קבוע חיובי. ניתן לראות כי}$$

לאחר מכן נעגל את מספר השרתים שחישבנו ע"פ נוסחאות (4.8)-(4.10) לשלם C_i תוך שמירה על -
 $C_1 + \dots + C_n = C$

דרך אחת פשוטה לעגל את גדלי קבוצות העבודה היא :

1. חישוב הערך הריאלי - $1 \leq i \leq n$, $C_i^* = \lfloor R_i \rfloor$.
2. חישוב $C^* = C_1^* + \dots + C_n^*$ (4.11) .
3. חישוב ההפרש $C - C^*$ (יש לשים לב ש- $0 \leq C - C^* \leq n-1$).
4. הדרך הפשוטה ביותר היא לשים את $C - C^*$, אחד בכל קבוצת עבודה על פי סדר יורד של $R_i - \lfloor R_i \rfloor$. אפשרות נוספת היא להוסיף לקבוצות העבודה לפי סדר יורד של $R_i - \lfloor R_i \rfloor / R_i$.

4.4 קביעת מטריצת מיומנויות השרתים A :

בהינתן גודל קבוצת עבודה, יש לקבוע ערכים התחלתיים למטריצה A. במקרה אמיתי קביעה זו צריכה להיות מבוססת על שילוב המיומנויות הקיים בפועל. בהנחה שמשתמשים באיגוד משאבים שלב זה לא אמור להיות קריטי במיוחד. בכל זאת, נגדיר פרוצדורה מסודרת לבניית A.

1. קביעת העמודה הראשונה ב-A כך ש- $\forall k, |G_k| = C_k$.
2. נסמן ב- $C_{i,j,k}$ את מספר השרתים בקבוצת עבודה i שיש להם רמת שירות j לשירות מסוג k.

$$(4.12) \quad C_{i,j,k} = \begin{cases} C_i & \text{for } j=1 \text{ and } k=i \\ \gamma_{i,j} (C - C_i)^{-1} C_k C_i & \text{otherwise} \end{cases}$$

כאשר $\gamma_{i,j}$ הוא פרופורציית השרתים בקבוצה i שיש להם מיומנות ברמה j.

לכל זוג (i,k) מתקיים $C_{i,j,k} \neq 0$ עבור ערך אחד בלבד של j (בקבוצה מסוימת לשרת יכולה להיות רמה אחת בלבד עבור כל סוג שירות).

אם $k=i$ ו- $j=1$ אזי מדובר במספר השרתים בעלי מיומנות i כמיומנות ראשית ולכן כל השרתים בקבוצה הם בעצם כל השרתים בקבוצה G_i ולכן $C_{i,j,k} = C_i$.

יש לשים לב שנוסחא 4.12 מביאה את מספר השרתים בקבוצה i שהם בעלי התמחות משנית k פרופורציונאלית גם לגודל קבוצה i - C_i וגם לגודל קבוצה k - C_k .

ניתן לראות כי מסגרת עבודה זו מאפשרת להשתמש גם במיומנויות חלקיות. לדוגמא – אם אנו רוצים שרק חצי מכל קבוצת עבודה יהיו בעלי מיומנות משנית, נבחר $\gamma_{i,2} = 1/2$. גם במקרה זה יש לעגל את מספר השרתים שמתקבל על מנת לקבל מספר שלם.

4.5 שימוש בסימולציה לקביעת שיפורים:

בהינתן זמני ההגעה, זמני השירות ומטריצה A כלשהי, אנו יכולים להשתמש בסימולציה על מנת להעריך את הביצועים. נמשיך לבצע שינויים על מנת למזער את מספר השרתים C ולמזער את K בהינתן C , תחת האילוצים שקבענו.

4.5.1 שלב האתחול של הסימולציה:

- אם המטרה אחוז הנענים בזמן מסוים לא תענה, אנו נגדיל את C באחד ונקטין את K באחד ונחזור על הניסוי. כך נשמור על $K+C$ קבוע. כאשר התהליך מוגדר, אנו נקבע קבוצות עבודה ומטריצה A . נחזור על צעד זה עד שנגיע למטרה של זמן ההמתנה. בגלל השימוש באיגוד משאבים, נחזה כי לא נצטרך להגדיל את C הרבה פעמים.
 - ברגע שמצאנו את רמת האיוש C שמביאה את מטרת זמן ההמתנה לאופטימום, נחקור את האילוף של אחוז הנחסמים. אם הגענו למטרה של אחוז הנחסמים, קיבלנו פתרון התחלתי פיזיבילי, ונוכל לעבור לשלב השיפורים. אם אנו לא עומדים באילוף זה, אנו צריכים להגדיל את K באחד ולבצע סימולציה נוספת, וכך הלאה עד שהאילוף של אחוז הנחסמים יתקיים. בעקבות תהליך זה יכול להיווצר מצב ש- C יהיה לא פיזיבילי. אם יש צורך בכך, נגדיל את C ביחידה אחת ונחזור על הסימולציה למציאת K פיזיבילי. בכך אנו מסיימים את שלב האתחול לחיפוש זוג (C, K) פיזיבילי, ומטריצה A פיזיבילית.
- לא נפסיק בשלב זה שכן יכולה להיות זוג טוב יותר. לדוגמא, העיגול שקובע את גודל כל קבוצת עבודה יכול לגרום לקשיים. לכן עכשיו נכניס שלב של ביצוע שיפורים. המטרה העיקרית שלנו היא להקטין את C , אך עתה נשתמש בביצועים הנצפים על מנת לזהות את השינוי המקומי המתאים.

4.5.2 ביצוע השיפורים:

- נקטין את מספר השרתים הכללי C ב-1 ונגדיל את מספר הקווים K ב-1, שוב תחת שמירה על $C+K$ קבוע. הרעיון הוא להוריד שרת מתוך השרתים בקבוצה בעלת הביצועים הטובים ביותר, כך שיש לו מיומנות משנית בקבוצה אחרת שגם לה ביצועים טובים. (קבוצה בעלת ביצועים טובים היא הקבוצה שמגיעה הכי קרוב למטרה של זמן המתנה).
- נעדכן את מטריצה A ע"י מחיקת השורה המתאימה ונחזור על הסימולציה. אם כל המטרות של זמן ההמתנה נענו, נחזור על הניסוי וננסה להוריד שרת נוסף.

- כאשר לא ניתן למחוק שרתים נוספים, ננסה לקחת את הזוג הלא פיזיבילי האחרון (C, K) וננסה לעשות אותו פיזיבילי, ע"י התאמה של שילוב כישורי השרתים. נחליף כישורים בין קבוצות בעלות ביצועים גבוהים לבין קבוצות בעלות ביצועים נמוכים.
- אם על ידי שינוי זה נצליח להביא את הזוג המדובר להיות פיזיבילי, נחזור לשלב הקודם וננסה להוריד את C ביחידה. כאשר שינוי שילובי המיומנויות נכשל נחזור לזוג הפיזיבילי האחרון ונעצור שם.

הרעיון הכללי הוא לבצע חיפוש מקומי של זוגות אפשריים ומטריצה A , ולהשתמש בסימולציה על מנת להעריך את הביצועים של כל זוג מוצלח. מכיוון שהביצועים הנדרשים לא מצריכים חילול בזמן אמת, ניתן להשתמש במחשב על מנת לחקור את כל השינויים המקומיים האפשריים אם נרצה בכך.

סיכום

המאמר מראה את נכונות הרעיון של *a little flexibility may go a long way* גם עבור מוקדי שירות טלפונים אשר פועלים במסגרת של ניתוב מבוסס מיומנויות. מסקנת המאמר היא שבמוקד כזה אין צורך בכך שכל השרתים יהיו אוניברסליים ובדוגמא של שש מיומנויות מספיק כי לכל שרת יהיו שתי מיומנויות בלבד על מנת להגיע לביצועים טובים כמו במקרה האוניברסלי.

פרק 5 –

– ניתוב מבוסס מיומנות –

סקירת המתרחש בפועל

5.1. ניתוב מבוסס מיומנויות במוקד השירות הטלפוני של בנק הפועלים

5.1.1 הקדמה -

מוקד "פועלים ישיר" נוסד ב-1995 כחלק מאסטרטגיית "בנק הפועלים" לספק ערוצי שירות מגוונים בנוסף לערוצים הקיימים (סניף ומכשירים בשירות עצמי) לכלל הלקוחות לפי בחירתם.

בשנת 1999 שודרג הערוץ הטלפוני ל- CONTACT CENTER, כאשר בנוסף לשדרוג התווספו המכירות היזומות (מכירה המתבצעת ע"י הוצאת שיחות ללקוחות הבנק) והמכירות הצולבות (מכירה המתבצעת לאחר שירות לקוח המתקשר לפועלים בטלפון). בנוסף הוקמו שירותי תמיכה ללקוחות האינטרנט.

בשנת 2002 החל הבנק בפרויקט שיפור המענה הטלפוני בסניפים, שכחלק ממנו הועברו שיחות הלקוחות מהסניפים ל"פועלים ישיר". מהלך זה הביא עמו את הקמת מרכז השירות השני, בנשר.

בשנת 2003 החל "פועלים בטלפון" ליישם מדיניות, הכוללת את עולמות התוכן הבאים:

- ערוץ שירות המשרת את כל המגעים הטלפוניים הנכנסים לחטיבה הקמעונאית.
- ערוץ מכירה המופעל ע"י החטיבה הקמעונאית, הן בשיחות נכנסות והן כמרכז טלמרקטינג.
- ערוץ תמיכה הן ללקוחות הבנק (אינטרנט פרטי, אינטרנט עסקי) והן לעובדי הבנק בנושאים טכניים ומקצועיים.
- ערוץ back office המטפל במטלות שאינן דורשות מגע ישיר עם הלקוח, ואשר נעשות כיום בסניפים. את הערוץ הנ"ל "פועלים בטלפון" מפעיל בשעות בהם לחץ השיחות יורד.

נתונים:

- קיימים כ- 500,000 לקוחות מנויים בפועלים ישיר, מתוכם 150,000 פעילים.
- שיעור הפעילות מכלל הבנק ב"פועלים ישיר" הינו 12%.
- קיימות כיום 150 עמדות בנשר ו-150 עמדות באזור.
- מספר הבנקאים בנשר עומד על 243, ומספר הבנקאים באזור עומד על כמאתיים.
- בפועלים ישיר מועסקים כשמונים אנשי מטה.

בנספח 1 מוצג המבנה הארגוני של "פועלים ישיר".

התאמה לצורכי הלקוח -

מלבד קו 2407*, הזמין 24 שעות ביממה, קיימים מספר קווים נוספים התואמים צרכים של סוגי לקוחות שונים ואשר לשירותם עומדים בנקאים עם כישורים מורחבים:

- קו תמיכה באינטרנט פרטי.
- קו תמיכה באינטרנט עסקי.
- קו תמיכה בלקוחות הבורסה.
- קו יועצים בעלי רישיון ייעוץ בשוק ההון.
- קו כתיבת אופציות במעו"ף.
- קו ני"ע זרים- בורסת ארה"ב.
- קו שירות בשפה הרוסית.
- קו שירות בשפה הערבית.
- קו לקוחות המתקשרים מחו"ל.
- קו ללקוחות חדשים הרוצים להצטרף לבנק הפועלים.
- קו שירות במשכן – קו זה משמש לשירות אנשים שלקחו משכנתא ורוצים לקבל מידע או רוצים לקחת משכנתא חדשה.

לכל אחד מסוגי השירות השונים ישנם נציגים בעלי מיומנויות מתאימות על מנת לענות ללקוחות המתקשרים לקווים השונים. כמו כן, לכל אחד מהמיומנויות השונות, מותאמת פעילות שלפיה משובצים הנציגים למשמרות הקיימות לפי הדרישה של כ"א בכל פעילות.

המיומנויות השונות של הנציגים הינן -

- | | |
|----------|----------------|
| • מכירות | • אינטרנט |
| • מעו"ף | • אינטרנט עסקי |
| • משכן | • שיפור המענה |
| • ני"ע | • יועץ |
| • ני"ע"ז | • לקוח חדש |
| • רוסית | |
| • שירות | |

שימור לקוחות קיימים וטיפול בלקוחות חדשים -

כחלק ממגוון השימושים בערוץ, "פועלים בטלפון" מספק את השירותים הבאים:

- שימור לקוחות קיימים - טיפול בלקוחות שזוהו כפוטנציאל לנטישה והבנק מעוניין לשמרם. שימור לקוחות קיימים מתבצע ע"י הצעת שדרוגים שונים ללקוחות הקיימים וכן הצעת מוצרים חדשים המתווספים לבנק.
- טיפול בלקוחות חדשים - יצירת קשר עם לקוחות חדשים, על מנת לבדוק את שביעות רצונם ולהגביר את נאמנותם לבנק. טיפול בלקוחות חדשים מתבצע ע"י הצעת מבצעים חדשים ואטרקציות חדשות.

שירות -

נושא השירות משמש כדגל בערוץ הטלפוני אשר השימוש החוזר בו ע"י לקוחות הבנק טמון בשלושת הנושאים הבאים:

- זמינות - 24 שעות ביממה ובזמני המתנה לפי סטנדרטים בינלאומיים (85% מן הלקוחות נענים תוך 25 שניות ו5% נטישה)
- מקצועיות - ידע וניסיון במגוון מוצרי הבנק, מתוך גישה שכל צורכי הבנק ימולאו בעת השיחה. על מנת לשמור על מקצועיות הבנקאים נדרשת רמת הכשרה גבוהה.
- שירותיות - אמינות, אדיבות ויוזמה.

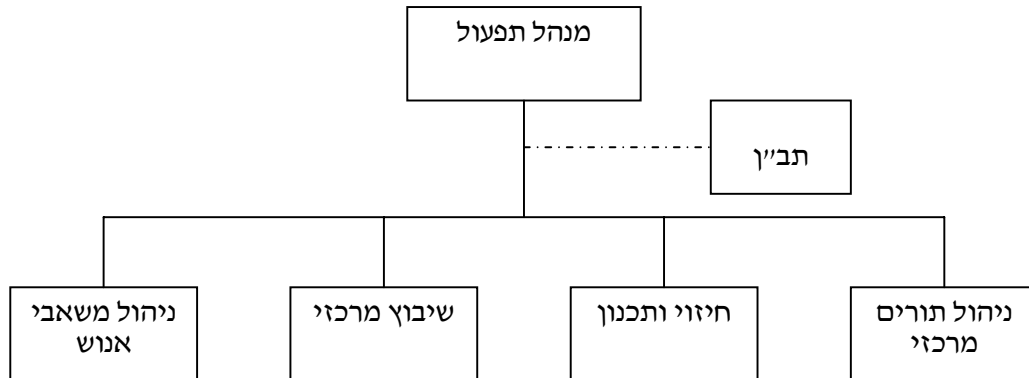
הסביבה -

הסביבה של "פועלים ישיר" מאופיינת ע"י:

- ריבוי אתרים.
- ארבעה פלחי לקוחות שונים.
- ארבעה עשר קווים שונים
- עשר סוגי פעילויות.
- מופע שיחות אקראי – כחמישים שיחות לדקה.
- אסטרטגיה לניתוב שיחות.
- מגוון רחב של מוצרים ומהליכים.

5.1.2 יחידת התפעול -

בתרשים 1 ניתן לראות את המבנה הארגוני של מחלקת התפעול:



תרשים 1

יחידת התפעול אחראית על החיזוי, התכנון, ניהול התורים ומשאבי אנוש.

- רפרנטית ניהול משאבי אנוש אחראית על כל התהליכים שקשורים בכ"א : גיוס, מיון כ"א חדש והנעת עובדים מצטיינים קדימה בשרשרת הניהולית.
- חיזוי ותכנון מתבצע הן ברמה השבועית והן ברמה החצי שנתית ע"י החזאי על מנת לתכנן נכון את המשאבים.
- שיבוץ מרכזי אחראי על בקרת המשבצים המקומיים בעבודתם, כלומר, לאחר שיבוץ הנציגים ע"י ה-WFM וביצוע שינויים ע"י הנציגים, מוצג למשבץ המרכזי השיבוץ הסופי והוא אחראי על כך שיהיו לכל פעילות מספיק נציגים לפי דרישת כ"א.
- ניהול תורים מרכזי הינו באחריות מנהל משמרת והם אחראים על ניהול נכון של המשמרת ע"י קבלת החלטות בזמן אמת כמו : הוצאת נציגים להדרכות/מכירות בזמני רוגע או החזרת נציגים לשירות בזמני עומס.

5.1.3 חיזוי ושיבוץ בטווח הקצר -

תהליך החיזוי מהווה את הבסיס לתכנון השימוש בכ"א, הן בטווח הקצר (שבוע) והן בטווח הארוך (בין חודש לשישה חודשים). חזאי הערוץ מכין חיזוי לטווח קצר ע"י שימוש בנתונים המתקבלים ישירות ממערכת ה-CTI תוך התחשבות בתבניות התנהגות של לקוחות, מצב שוק ההון ותהליכים אחרים בהם שותף הערוץ כגון שיפור המענה הטלפוני המוסיף שיחות לערוץ הטלפוני.

על בסיס חיזוי זה, מכין החזאי דרישה לשני המוקדים. גרף דרישה זה כולל את תחזית ההיעדרויות ואת הצרכים לשירות שיחות נכנסות, שיחות לקווים המיוחדים, מכירות, תדרכים, הדרכות ו- back office. גרף הדרישה מוזן למערכת השיבוץ-ה-WFM. גורם נוסף המוזרם למערכת ה-WFM הינו אילוצי העובדים. נתונים אלו מוזרמים ע"י משבצים מקומיים בהתאם למדיניות הערוץ, כאשר ניתנת קדימות ללימודים מחד ולצרכי הורים מאידך. מערכת ה-WFM מייצרת שיבוץ לשבוע הקרוב, ולעובדים יש יומיים להגיב ולבקש שינויים לפי שהשיבוץ הופך להיות פורמאלי. מערכת השיבוץ משבצת הפסקות לפי כללים שהוזרמו ע"י החזאי והנהלת הערוץ. ככלל, כל עובד מקבל הפסקה כל שעותיים.

גורמים המשפיעים על חיזוי שיחות -

ישנם מספר נתונים כמותיים המשפיעים על חיזוי השיחות:

- נתונים היסטוריים.
- תאריך או יום בשבוע.
- ותקלות במערכת.

בנוסף, קיימים גם נתונים איכותיים המשפיעים על החיזוי:

- עונתיות.
- שינוי/הוספת פעילויות.
- ומצב מדינאי/כלכלי.

גורמים נוספים המשפיעים על חיזוי השיחות הנם פוטנציאל פעילות הכולל:

- אחוז גידול טבעי בפעילויות.
- הוספת גלי שיפור המענה הטלפוני.
- הצטרפות פעילויות חדשות (לדוגמא - לקוח חדש או לקוחות משכן).
- צרכי הדרכה.
- צרכי back office.

תכנון לטווח הארוך -

תכנון לטווח הארוך מבוצע ע"י החזאי. תכנון זה לוקח בחשבון את הגורמים הבאים:

- המצב הכלכלי הצפוי והשפעות על שוק ההון
- מוצרים ושירותים חדשים הצפויים להיכנס בתקופת התכנון, הן בערוץ עצמו והן במקומות אחרים בבנק.
- הוראות בנק ישראל כמו גילוי נאות, המוסיף זמן לשיחה.
- מבצעי מכירה של החטיבה.

שלבי התכנון -

תרשים 2 מציג את שלבי התכנון של המערכת בטווח הארוך :

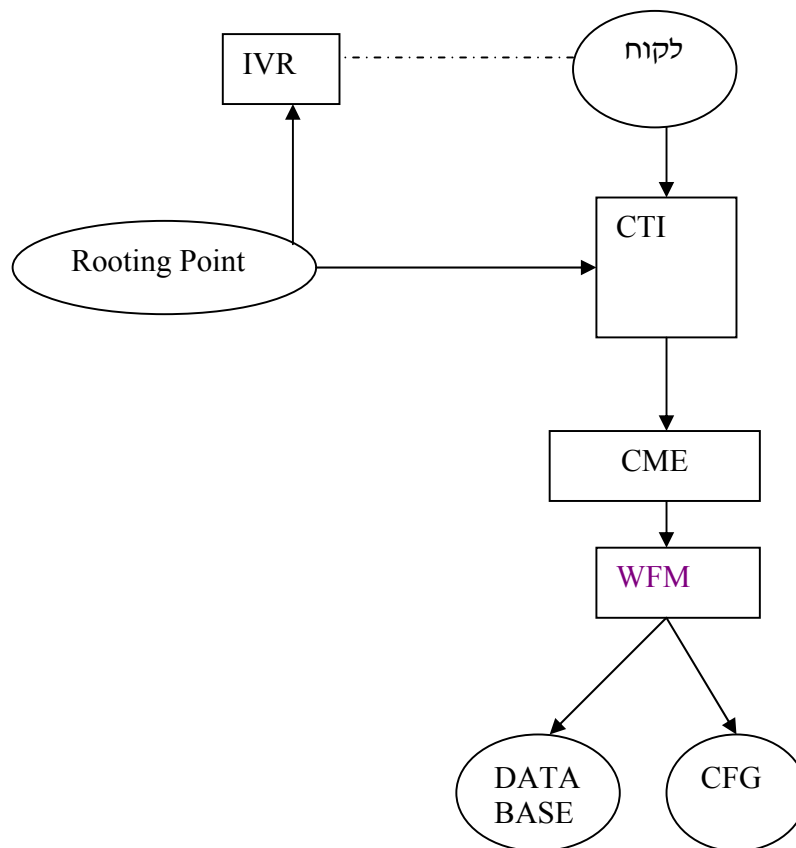


תרשים 2

5.1.4 מערכת ה-WFM -

מערכת ה-WFM משמשת לשיבוץ עובדים למשמרות בהתחשב ב :

- סוגי העובדים.
 - סוגי המשמרות.
 - סוגי הפעילויות במוקד.
 - אילוצי העובדים.
 - הפסקות.
 - גרפי דרישה לכל אחת מן הפעילויות הקיימות במוקד.
- מערכת ה-WFM נותנת מענה למורכבות הנובעת מפיצול המוקדים, ריבוי פעילויות ובקשות עובדים.
- בתרשים 3** מוצג אפיון הקשר בין המערכות הטכנולוגיות למערכת ה-WFM .



תרשים 3

כל שיחה של לקוח מתחילה בהגעה למרכזיה של ה-CTI שהיא מערכת המבצעת אינטגרציה בין טלפונים למחשבים, כלומר שליטה על הטלפונים באמצעות המחשב. במידה והלקוח רוצה להגיע למענה הקולי על מנת לשמוע מידע ממוחשב על פרטי החשבון שלו, אזי הטיפול בשיחה נפסק ומועבר ל-IVR.

במידה והלקוח רוצה לדבר עם נציג שירות, אזי ה-CTI מעבירה פקודה למרכזיה לאיזה נציג להעביר את השיחה. ה-CTI מנהלת סוגי פניות שונות, אל ומהארגון, של לקוחות למול המשאבים המיועדים לטיפול בפניות אלו. במידה והלקוח רוצה לאחר שמיעת פרטים במענה הקולי לדבר עם נציג שירות, הוא עובר דרך צומת ניתוב (routing point) ל-CTI שמקבל את ה-"פיקוד" על השיחה. ב-ROUTING POINT "חונות" השיחות הנכנסות לפני ניתוב לגורם המטפל.

במסגרת הטיפול בפניה ע"י ה-CTI מתקיימים התהליכים הבאים :

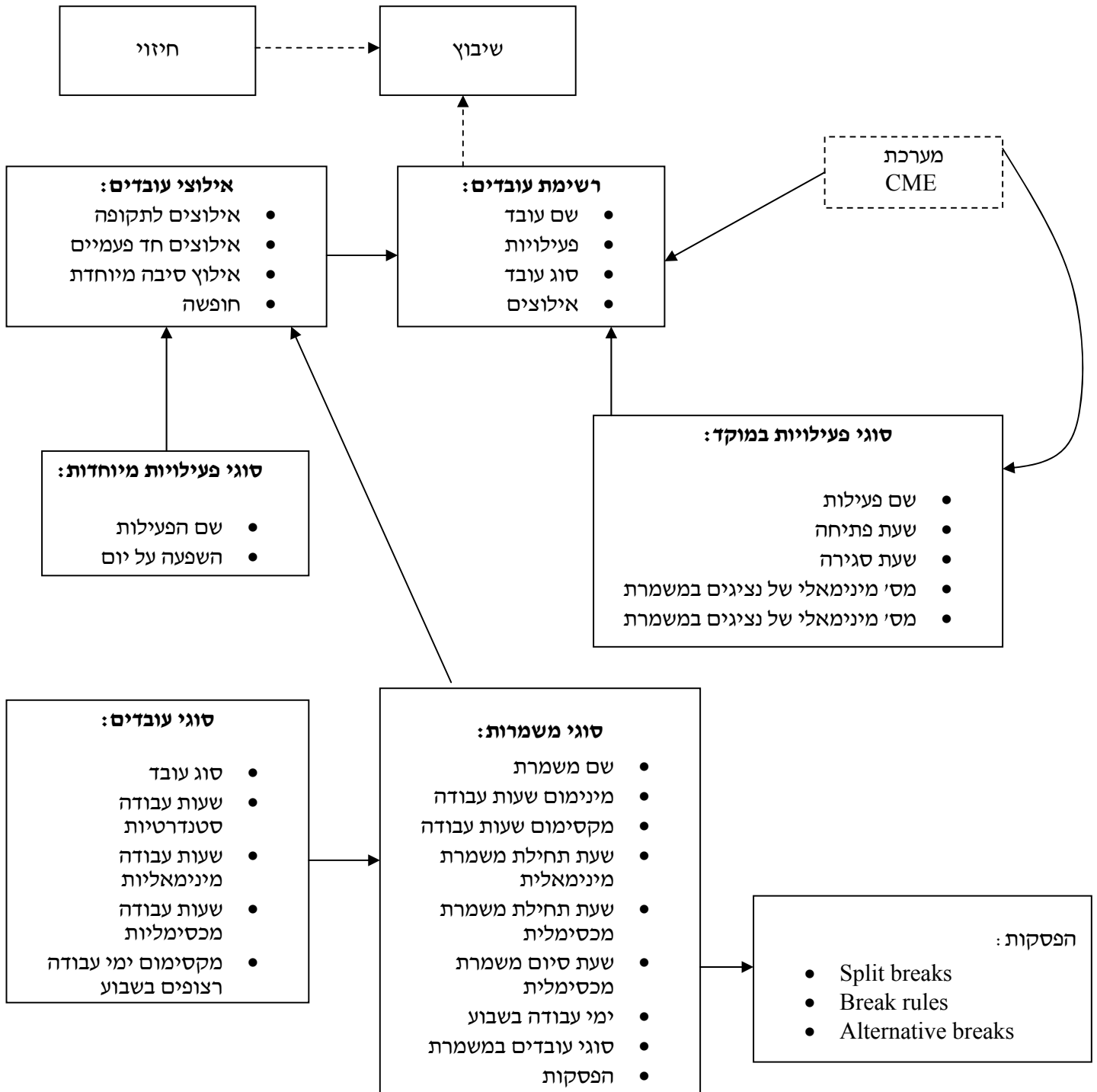
- הגדרה ומיון של הפניות.
- קבלת החלטה לגבי דרך הטיפול הנדרשת.
- ניהול תורי הפניות השונות הקיימות במערכת.
- ניתוב השיחה למשאב הזמין לפי הגדרת הטיפול המתאימה לפניה.
- ניהול והגדרה של המשאבים המיועדים לטיפול בפניות.
- מעקב אחר ביצועים.
- חיזוי ותכנון כמות המשאבים הנדרשת.

לדוגמא , ב- CTI נבדקים הנתונים של הלקוח כמו : סוג לקוח (רגיל או פרטי) וכן סוג ההתמחות המתאימה לו שלה נציג פנוי כרגע. לאחר ה-CTI מועברת השיחה ל- CME, שמכילה את בסיס הנתונים של כל הנציגים לפי מספר עובד. בסיס הנתונים מכיל למעשה את אוסף כל הנתונים המוגדרים ע"י הארגון או המערכת, ומצורפים לפניה לצורך הגדרה ומיון למול יתר הפניות הקיימות במערכת. הנתונים המצורפים הינם מספר עובד, המיומנויות שלו ורמת ההתמחות שלו בכל אחת מן המיומנויות. על סמך נתונים אלו המרכזייה מחליטה על העברת השיחה לנציג המתאים לו. ה- CME הינו תור פניות המנוהל במרכזיה מול קבוצת נציגים מוגדרת (ללא יכולת הפעלה של עדיפות, אלא ברמה של FCFS).

מערכת ה-WFM שהינה מערכת השיבוץ מכילה את נתוני העובדים לפי שמם וההתמחות שלהם. נתוני העובדים נמצאים ב-CFG. ב- DATA BASE נמצאים נתוני המדיניות והקונפיגורציות של מוקד השירות, כלומר, אילו משמרות קיימות, אלו מיומנויות קיימות, אלו אילוצי וסוגי עובדים קיימים ואלו סוגי פעילויות קיימות. ה-WFM מראה נתונים בפועל של כל עובד ועובד במערכת בפרקי זמן של רבע שעה.

כיום קיימות ב"פועלים ישיר" שתי מרכזיות נפרדות , מכיוון שמבחינה גיאוגרפית קיימים שני מוקדים שונים. לכן ברמת השיבוץ ה-WFM משבץ שני שיבוצים נפרדים. אך ברמת המענה בפועל כל שיחה שנכנסת יכולה להיכנס לאחת מן שתי המרכזיות באופן אקראי.

בתרשים 4 מוצג אפיון ע"י טבלאות הינו הקשר בין המדיניות, הקונפיגורציה והעובדים במערכת ה-WFM:



תרשים 4

שלב 1 –

בניית גרף חיזוי שבועי לכל אחת מהפעילויות .

גרף החיזוי נבנה עפ"י נתונים כמותיים מתאריכים קודמים, לדוגמא, אם ברצוננו לחזות כמה שיחות יתקבלו ב-10 לחודש מסתכלים על ה-10 לחודש של חודשים קודמים או תאריכים קרובים ל-10 לחודש כמו, ה-9, 8 או ה-11. תאריכים קרובים בחודשים קרובים לא מספיקים על מנת לחזות את כמות השיחות החזויה כי היום בשבוע גם כן חשוב. לדוגמא- אם ה-10 לחודש נופל על יום ראשון ושאר התאריכים הקרובים ל-10 לחודש בחודשים הקודמים נופלים על ימים שאינם ימי ראשון, אזי ימים אלו אינם טובים בשביל להסתמך עליהם לחיזוי. ולכן, על מנת לחזות באופן המיטבי את כמות הנציגים הרצויה לנו הבנק מסתמך גם על תאריך בחודש וגם על יום בשבוע. אלמנטים נוספים המשפיעים על השיבוץ הינם חגים, תקלות או כל דבר שאינו שגרתי בתאריכים מסוימים. לדוגמא, לאחר פסח היה עומס רב של שיחות עקב החזרה מהחגים, ולכן לא ניתן להסתמך על תאריכים אלו בבניית חיזוי שבועי.

דוגמא נוספת הינה תקלה ביום מסוים שגרמה לאי קבלת שיחות בשעות מסוימות והיגררות השיחות לשעות מאוחרות יותר ובכך נגרם אי סדר מבחינת העומס היומי בשיחות. גורמי השפעה נוספים על החיזוי הינם גורמים איכותיים כמו: עונתיות, שינויים במערכת, שינויים בפעילות ומצב כלכלי או מדיני. דוגמא לעונתיות הינה ההבדל בשיחות בזמן החופש הגדול, שאז כמות השיחות יחסית קטנה מכיוון שמרבית הלקוחות נמצאים בחופשות משפחתיות. דוגמאות לשינויים במערכת ע"י גורמים איכותיים:

- הכנסת שינוי אשר עוזר לנציגים לענות ללקוחות בפרק זמן קצר יותר ומוריד את זמן השיחה.
- דוגמא לשינוי בפעילות הינה הכנסת שיפור במענה הטלפוני שבעצם הסיט שיחות מן הסניפים ל"פועלים ישיר" ובכך גרם להגדלת כמות השיחות.
- דוגמא לשינוי במצב מדיני/כלכלי הינו לדוגמא חיסול אחד מבכירי החמאס או פיגוע התאבדות.

בנוסף להתחשבות בגורמים כמותיים וגורמים איכותיים לחיזוי השיחות ישנה גם התחשבות בפוטנציאל פעילות שכולל צרכי הדרכה או מכירות לאותו השבוע, צרכי ה- back office לאותו שבוע וכן התחשבות באחוז גידול טבעי של השימוש בשירות.

שלב 2 –

על סמך שיחות חזויות לכל שעה נתונה מוצאים את ה-AHT, שזהו זמן טיפול ממוצע לשיחות בשעה. ה-AHT מחושב ע"י מכפלת מספר השיחות החזויות לשעה כפול זמן טיפול ממוצע לשיחה. זמן טיפול ממוצע לשיחה מורכב משני מרכיבים: האחד זמן שיחה ממוצע עם לקוח והשני זמן wrap-up - מילוי טופס, עדכון נתוני לקוח וכו'.

בדיקת ה-AHT נותן עוד מדד לבחינת ההשתנות היום-יומית של זמני השיחה למול שעות היום, הימים בשבוע וכו' וגורמת לנו לבצע החלטות יותר נכונות בחיזוי העתידי.

שלב 3 -

לאחר בניית גרף חיזוי לכל פעילות יש למצוא לפי נוסחת $ERLANG C$ את מספר הנציגים הדרושים בכל שעה ושעה. בנוסחת $ERLANG C$ מוכנסים הפרמטרים הבאים:

- זמן שיחה ממוצע.
- אחוז שיחות שיענו תוך זמן דרוש.
- זמן המתנה מקסימאלי עד למענה.

נוסחת $ERLANG C$ לוקחת בחשבון את אקראיות ההגעה, ההמתנה בתור עד לקבלת שירות וכן כמות לקוחות אינסופית, ומחשבת על סמך $PASTA$ את $P\{wait>0\}$, כלומר – ההסתברות להמתנה.

שלב 4 -

לאחר הכנסת נתונים למערכת, בוחרים את מספר הנציגים הדרוש לנו לפי נצילות המוקדנים וכן לפי אחוז הנטישות שמתקבל. העומס הרצוי על המוקדנים הינו 83% לשעה כיוון שצריך לקחת בחשבון הפסקות. אחוז הנטישה שאנו לוקחים בפועל לא עולה על 5% לפי האסטרטגיה של המוקד. כמו כן, נדרש כי 80% מהלקוחות ייענו בתוך 20 שניות.

שלב 5 -

לאחר בחירת מספר המוקדנים הרצויים לכל פעילות בכל שעה ושעה מחלקים את המוקדנים הרצויים בין שני המוקדים לפי יחס הנציגים הקיימים בפועל. מערכת ה- WFM מפזרת את השיבוצים (משמרות ושעות) באופן שוויוני בין כל הבנקאים.

שלב 6 -

השלב הבא הינו הכנסת נתונים נוספים הרלבנטים לשיבוץ. נתונים אלו נקבעים לפי אסטרטגית המדיניות במוקד. כל שיבוץ מורכב ממשמרות שלהן שעת התחלה ושעת סיום וכן מימי עבודה אפשריים לפעילות. לכל משמרת מתאימים את סוגי העובדים המתאימים למשמרת וכן בוחרים את ההפסקות המתאימות. כמו כן, כל שיבוץ מורכב מנציגים השייכים לראש צוות מסוים. לכל נציג מגדירים סוג עובד בכדי שניתן יהיה לשבץ אותו בהתאם לחוזה העבודה שלו. בנוסף מגדירים לכל נציג את המיומנויות השייכות לו ועל פיהם הנציג ישובץ לפעילויות השונות במוקד. לכל פעילות יש מיומנות מסויימת שבד"כ מיוצגת ע"י שם הפעילות. בנוסף, לכל נציג יש תבנית אילוצים קבועה משלו שעל פיה הוא משובץ במשך השבוע. התבנית מוגדרת מאילוצים קבועים שלא ניתנים לשינוי ומאילוצים לא קבועים שכן ניתנים לשינוי. כלומר, אם נציג מסוים סימן את יום רביעי כיום שהוא רק יכול לעבוד בבוקר כחובה כי בערב יש לו לימודים הוא לא ישובץ אחר הצהריים, אך אם אין לאותו עובד ביום רביעי אילוץ כלשהוא הוא ישובץ בהתאם לאילוצי המערכת באותו יום. כלומר, אם המערכת תצטרך אותו ביום רביעי בבוקר הוא ישובץ בבוקר וכנ"ל לגבי הערב.

לאחר הכנסת כל נתוני המדיניות של המערכת, נתוני העובדים ואילווציהם וכן כל גרפי הדרישה לכל הפעילויות יש צורך בקביעת עדיפות שעל פיה המערכת תשבץ את העובדים. כלומר אם יודעים שרוב השיחות שננטשות שייכות לפעילות מסוימת, תינתן עדיפות לאותה פעילות ע"י קביעת רמת שירות מקסימאלית. כמו כן, המערכת משבצת נציגים בהתאם למדיניות וכן בהתאם לחוסרים הקיימים על מנת למלא את גרף הדרישה הקיים לכל פעילות. פרמטרים נוספים שאנו מכניסים לכל פעילות על מנת שתשבץ בצורה המקסימלית את הנציגים הם :

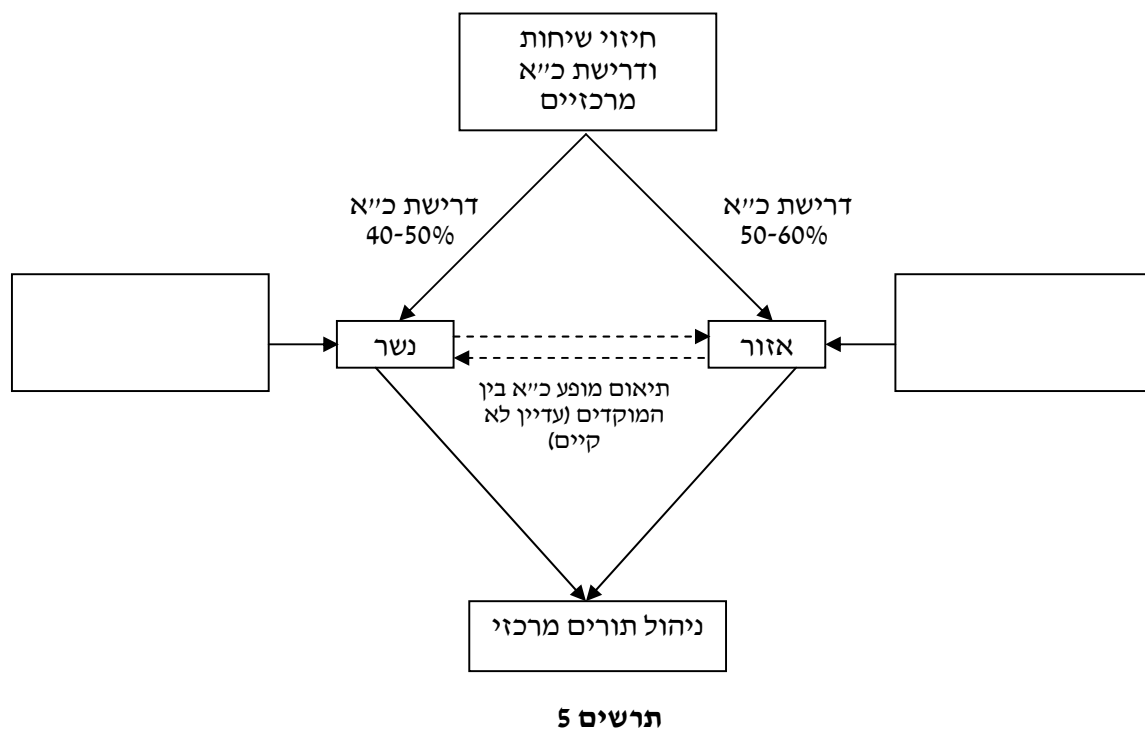
- אחוז עבודה לא פרודוקטיבית.
- עומס לא כולל הפסקות.
- אחוז הפסקות.
- הדרכות.
- אחוז היעדרות.
- זמן מהירות תגובה.
- אחוז תפוסה.
- לאחר כמה זמן שיחות נחשבות לננטשות.
- אחוז שיחות ננטשות.

לאחר כל שבוע מתבצע משוב הבא להשוות את ההבדל בין כמות השיחות בפועל לכמות השיחות החזויה. לאחר זיהוי הפערים מנסים למצוא סיבות לפערים אלו כמו : תקלות, אירועים חריגים וכו'. כמו כן, מתבצע משוב של השוואת כמות העובדים בפועל למול השיבוץ וגרף הדרישה. גם כאן מנסים לאחר זיהוי הפערים למצוא סיבות לפערים אלו כמו : היעדרות גבוהה וכו'.

5.1.5 ניהול תורים -

"פועלים בטלפון" מנהל סט תורים משותף לשני האתרים שבו נמצאים הבנקאים. באופן פעולה זה, בא לידי ביטוי המוקד הווירטואלי שבו נכנסות שיחות הן לאזור והן לנשר, ולאחר הזדהות במענה הקולי וקבלת מידע, מועברת השיחה לבנקאי זמין באחד משני האתרים. מנהלי התורים אחראים על רמת שירות, ומשתמשים בכלים העומדים לרשותם על מנת לקבל החלטות שונות, כמו הוספה או הורדה של בנקאים. אחריות נוספת של מנהלי המשמרת היא להחליט הקצאת משאבים מתאימים לכמות השיחות המגיעה למוקדים. כלומר, במידה ויש בנקאים פנויים לאורך זמן צריכה להתקבל החלטה לשחרר בנקאים למשימות אחרות ובמידה וקיימות שיחות ממתינות. באחריות מנהלי משמרת לדרוש בנקאים ממנהלי המוקדים. תקשורת מנהלי התורים נעשית אך ורק בתאום עם מנהלי המוקדים ואו סגניהם ולא מול מנהלי הצוותים או הבנקאים. מנהלי המוקדים ואו סגניהם אחראים לספק את העובדים הנדרשים בתוך 10 דקות ואו לקלוט עובדים לצורך מטלות כגון: מכירות הדרכות, back office וכו'.

בתרשים 5 מוצג הקשר בין המשבץ המרכזי, המשבצים המקומיים ומנהלי התורים -

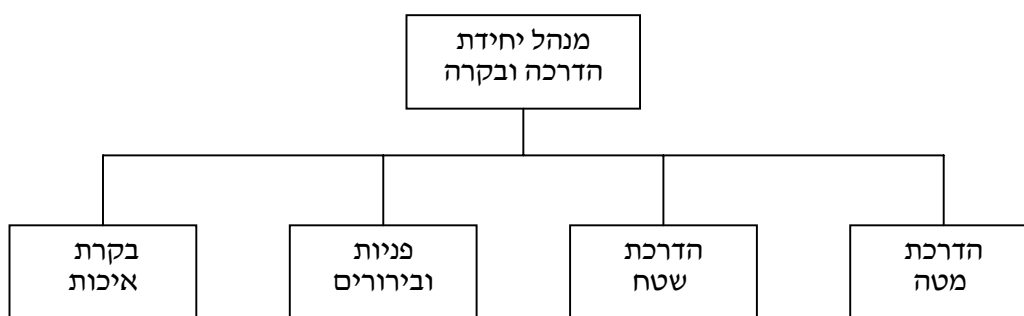


5.1.6 יחידת ההדרכה והבקרה -

יחידת ההדרכה והבקרה הינה אחראית על נושאים שונים :

- הדרכת נציגים בקורסים השונים לפני התחלת עבודתם במוקד ע"י מחלקת הדרכה-שטח.
- הדרכת הנציגים בזמן עבודתם במוקד ע"י מחלקת הדרכה מטה.
- קבלת פניות וברורים מלקוחות לא מרוצים וטיפול בפניות אלו.
- ביצוע בקורות איכות לכל תקופה על מנת לבדוק את רמת השירותיות והמקצועיות של כל נציג.

בתרשים 6 מוצג המבנה הארגוני של היחידה :



תרשים 6

בקורות איכות, הדרכה ותגמול -

- כל עובד שעובד במשמרת של 8 שעות מקבל שעה הפסקה ו- 20 דקות תדריך לפני תחילת המשמרת.
- כל עובד מקבל במשך החודש 7 שעות של חזיוקים והעשרות.
- אחת לרבעון מקבל העובד משוב מבקרת איכות המתבצעת ע"י יחידת בקרת איכות. המשוב ניתן על סמך האזנה ל- 15 שיחות אקראיות של הבנקאי. במשוב מבצעים הפקת לקחים ובניית תוכנית לשיפור.
- העובד מקבל בנוסף משוב ממנהל צוותו על סמך האזנה לשיחותיו במשך החודש.
- התגמול לעובד מורכב משכר בסיס, ותגמול תקופתי אשר מושפע מתוצאות בקורות האיכות של הבנקאי. וממכירות.

מערך ההדרכה -

מערכת ההדרכה כולל בתוכה הדרכת מוצרים/הליכים חדשים, שינויים במוצרים/הליכים קיימים, שינויים בנהלים, ריביות, תנאי שירות לפלחי לקוחות, חידושים ושינויים בשוק ההון ומבצעים.

טיפול בפניות לקוחות -

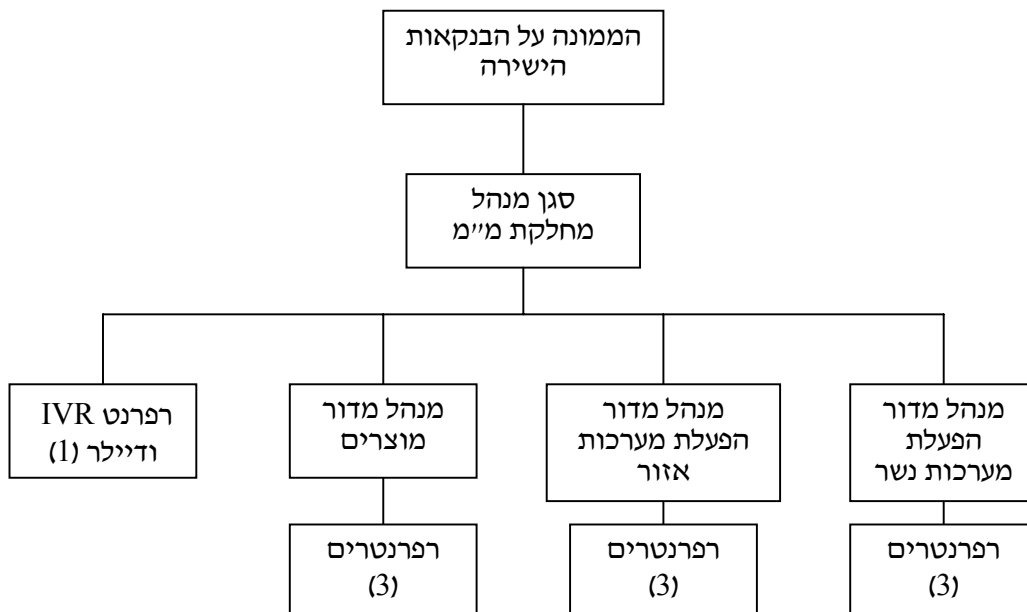
למערכת מגיעות בכל חודש 1,000 פניות של לקוחות בממוצע. סוגי הפניות השונות קשורות בטעויות של הנציגים ע"י מתן תשובה או הסבר לא נכון ללקוח, או לחילופין קשורה לא בטעות שנעשתה ע"י הנציג, אך הבנה לא נכונה של הלקוח. כ-70% מהפניות מטופלות ע"י ראש הצוות באותו היום וכ-30% מתוכן מטופלות ע"י צוות הפניות תוך 3 ימים. תהליך הטיפול בפנייה כולל:

- קבלת הפנייה באמצעות הטלפון, מכתב או ע"י הסניף.
- האזנה לשיחה שנתקבלה.
- קבלת החלטה מקצועית.
- טיפול מול גורם רלבנטי.
- תיקון מול הסניף.
- תיעוד והפקת לקחים.
- משוב לנציג.
- פרסום פנימי של הפנייה לכל הנציגים האחרים.

5.1.7 יחידת תפעול מערכות -

יחידת תפעול מערכות אחראית על ניהול תקין של כל מערכות המידע השונות הקיימות בשני מוקדי השירות. מעל שני המוקדים ממונה אחראי אשר בתפקידו לטפל בבעיות שונות הנוצרות במערכת, ולכל מוקד ממונה אחראי מערכות. מתחת לכל אחראי מערכות מידע במוקדים עובדים כשלושה רפרנטים האחראים על כל המחשבים ומערכות השונות של המוקד. כמו כן קיים מנהל מדור מוצרים שאחראי על המוצרים השונים שמתופעלים ע"י מערכות המידע, כמו בניית דוחות שונים, בניית מערכות כ"א וכו'. לאחראי על מדור מוצרים ישנם שלושה רפרנטים העוזרים לו בכך. כמו כן, קיים רפרנט IVR האחראי על ניהול נכון של המענה הקולי וטיפול בבעיות שונות הקשורות אליו.

בתרשים 7 מובא המבנה הארגוני של יחידת תפעול המערכות:



תרשים 7

ניהול תורים ע"י מערכות המידע -

ניהול התורים ע"י מערכות המידע מתחיל בכניסת שיחה למרכזיית ה-CTI, שבה מנהלים סוגי פניות שונות של הלקוחות למול המשאבים המיועדים לטיפול בפניות אלו. במסגרת הטיפול מתקיימים התהליכים הבאים:

- הגדרה ומיון של הפניות.
 - קבלת החלטה לגבי דרך הטיפול הנדרשת.
 - ניהול תורי הפניות השונות הקיימות במערכת.
 - ניתוב השיחה למשאב הזמין לפי הגדרת הטיפול המתאים לפניה.
 - ניהול והגדרה של המשאבים המיועדים לטיפול בפניות.
 - מעקב אחר ביצועים.
 - חיזוי ותכנון כמות המשאבים הנדרשת.
- ה-CTI מעבירה את השיחה לנציג לפי אסטרטגיה המגדירה את אופן הטיפול הפנייה וכן את אסף ה-DATA BASE המצורפים לפניה לצורך הגדרה ומיון למול יתר הפניות הקיימות במערכת.
- ניהול תורי הפניות מבוצע באמצעות הרכיבים הבאים:

- **ROUTING POINT** - רכיב לוגי במרכזיה בו "חונות" השיחות הנכנסות לפני ניתוב לגורם המטפל
- **VIRTUAL QUEUE** - רכיב לוגי המתקיים במערכת ומאפשר לחלק את ה-ROUTING POINT למספר תורי משנה (מופעל בעיקר לצורך טיפול בפניות מסוג פקס ו-EMAIL).
- **ACD** - תור פניות המנוהל במרכזייה מולקבוצת נציגים מוגדרת (ללא יכולת הפעלה של עדיפות, אלא ברמה של FIFO). ב-ACD מוגדרים אוסף הנתונים אותם רוצה המערכת למדוד.

אסטרטגיית ניהול תורים -

תהליך ניהול התורים הינו התהליך לפיו נקבע ניתוב הלקוח לנציג שירות לפי סוג לקוח , פילוח הלקוח, זמן המתנה, היסטוריית שימוש ונציגים פנויים בעלי מיומנות מתאימה לטפל בלקוח.

לכל פעילות קיימת מיומנות מתאימה, אך לא באופן חד חד ערכי. כלומר - יכולה להיות אופציה שאותה המיומנות תהיה שייכת לכמה פעילויות בו זמנית. לכל מיומנות ישנה דרגת התמחות. ככל שרמת ההתמחות של הנציג יותר גבוהה, כך הנציג יהיה זמין יותר ללקוחות מועדפים. רמת התמחות נקבעת עבור כל נציג. ככל שנציג יותר ותיק ובעל רמת שירותיות ומקצועיות גבוהה, כך רמת ההתמחות שתקבע לכל אחת מן המיומנויות תהיה גבוהה יותר.

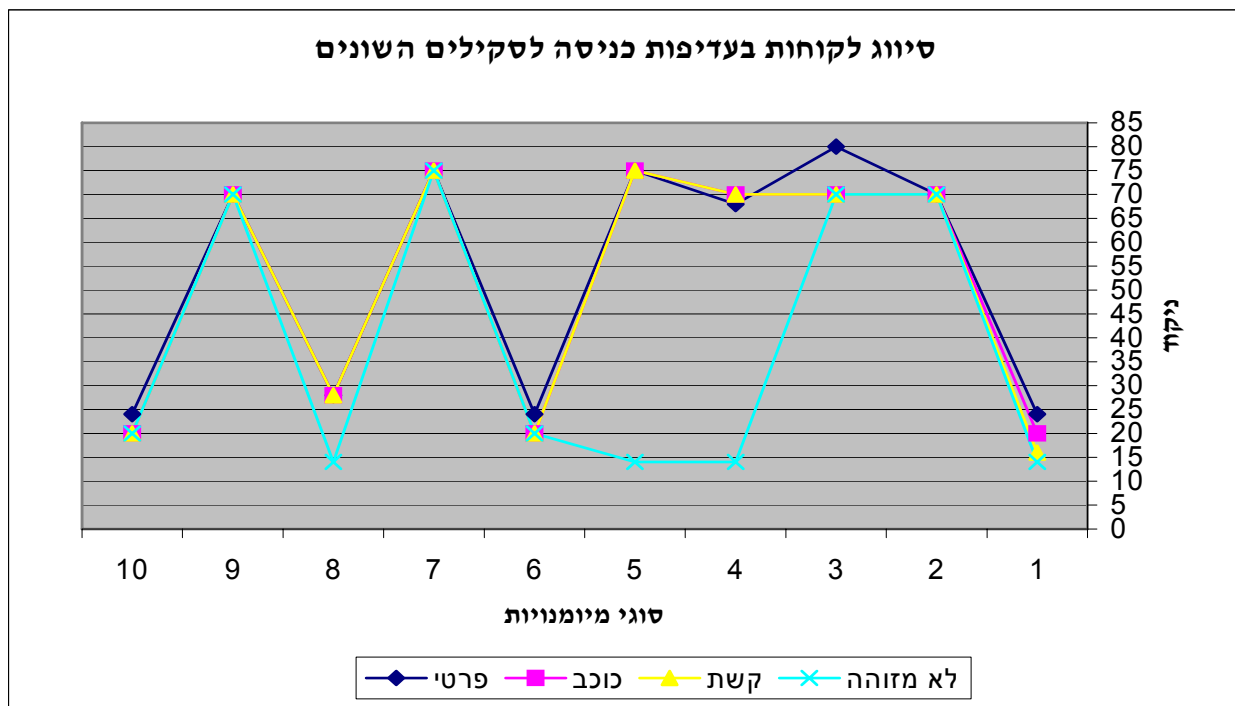
כל לקוח מפולג ל- 4 סוגי לקוחות: קשת, כוכב, פרטי ולא מזוהה. חלוקה זו קיימת לכל אחד מן הקווים הזמינים ללקוח - כלומר קו שירות, הסטה, ניירות ערך זרים, אינטרנט, אינטרנט עסקי, רוסית, משכן, מסילה ויועץ. כאשר לקוח מתקשר לקו מסוים הוא נכנס להמתנה לנציגים בעלי אותה מיומנות השייכת לפעילות אליו התקשר הלקוח.

לכל אחד מן הקווים יש צומת ניתוב (routing point) בה הלקוח מנותב לפי מספר קריטריונים:

הקריטריון הראשוני הינו עדיפות כניסה איתה מגיע כל לקוח למערכת. כלומר, למשל ללקוח פרטי תהייה עדיפות כניסה גבוהה יותר מלקוח קשת וכוכב. כמו כן, עבור הקווים המקצועיים עדיפות הכניסה הינה גבוהה יותר מאשר עבור הקווים הרגילים.

לאחר עדיפות הכניסה הראשונית מועלה עדיפותו של כל לקוח בניקוד מסוים כעבור 20 שניות. במידה ותוך דקה עדיין הלקוח לא קיבל שירות, עדיפותו ממשיכה לגדול על פני אחרים. העדיפות ממשיכה לגדול עבור כל דקה נוספת שהלקוח ממתין ולא קיבל שירות

בגרף 1 מוצג הניתוח שנעשה עבור הקווים השונים וסוגי הלקוחות השונים:



גרף 1

כאשר המקרא למיומנויות השונות מוצג בטבלה 1:

שירות	1
אינטרנט	2
אינטרנט עסקי	3
ניע"ז	4
מעו"ף	5
הסטה	6
מסילה	7
ניע	8
יועץ	9
רוסית	10

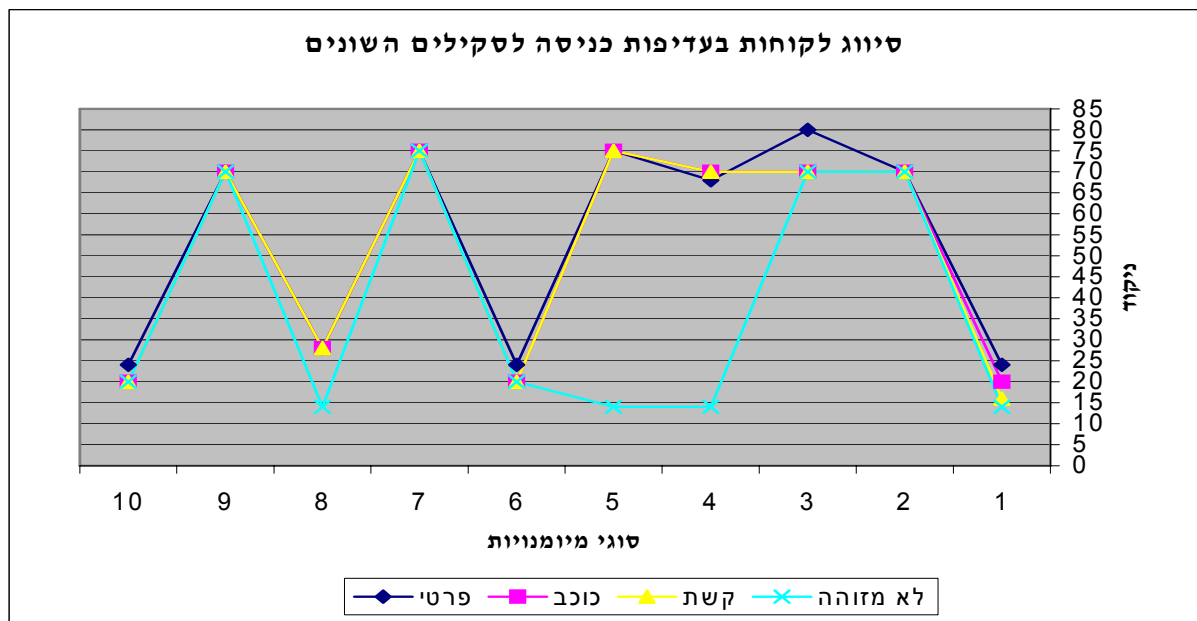
טבלה 1

ניתן לראות כי המיומנות בעלת עדיפות הכניסה הגבוהה ביותר היא אינטרנט עסקי ולאחריה מסילה. לאחר 20 שניות העדיפות משתנה ומעו"ף ומסילה מקבלות את הניקוד הגבוה ביותר ואחריהן אינטרנט, אינטרנט עסקי וניע"ז. לאחר דקה של המתנה בתור העדיפות העליונה לקבלת שירות ניתנת למסילה ומעו"ף ולאחריהן אינטרנט, אינטרנט עסקי, יועץ וניע"ז.

ניתן לראות כי בכל העדיפויות הקיימות (כניסה, לאחר 20 שניות ולאחר דקה) המיומנויות המקצועיות הנן בעלות עדיפות כניסה גבוהה מן מיומנויות השירות, ההסטה וני"ע. הסיבה לכך הינה אסטרטגיה של הבנקאות הטלפונית לפיה מיומנויות מקצועיות עדיפות על פני מיומנויות שירותיות, מכיוון שלקוחות המשתמשים בקווים מקצועיים יותר רווחיים לבנק.

בגרף 2 מוצג הניקוד עבור שאר סוגי הלקוחות (קשת, כוכב ולא מזוהה)

בגרף ניתן לראות את הדירוג בעדיפות לפי סוג לקוח. העדיפות הינה: לקוח פרטי, לקוח כוכב, לקוח קשת ולקוח לא מזוהה. כמו כן, ניתן לראות כי עדיפות הכניסה למיומנויות המקצועיות גבוהה מזו של מיומנויות השירות, ההסטה וני"ע.



גרף 2

- מקרא המיומנויות מוצג בטבלה 1 לעיל.
- התרשים הנ"ל נערך עבור עדיפות הכניסה, אך ניתן ליצור את אותו תרשים למול עדיפות לאחר 20 שניות ועדיפות לאחר דקה.

אופן טיפול בשיחות נכנסות -

הפרמטרים להגדרת פנייה הם :

- פלח- הגדרה עסקית למאפייני הלקוח.
- סוג שירות- מהות הפניה לארגון (מזוהה לפי בחירת הלקוח ב-IVR, היסטוריית פניות, מספר הטלפון אליו חייג הלקוח).

אוסף הפרמטרים הנ"ל יוצר פונקציה בעלת שני משתנים : פלח ושירות. פונקציה זו בעלת טווח ערכים רחב וסופי, כאשר כל תוצאה בטווח מקבלת התייחסות תהליכית של טיפול.

תהליך הטיפול בפנייה נכנסת -

1. קליטת השיחה ב-IVR וטיפול לפי סוג השירות (מספר הטלפון המחוייג ע"י הלקוח).
2. בקשת ניתוב לנציג מה-IVR ל-ROUTING POINT המכיל אסטרטגיה לגבי תהליך הטיפול בפנייה. אסטרטגיית הניתוב לגבי כל אחד מסוגי השירות מכילה הגדרות של עדיפות ראשונית (כניסה להמתנה), הגדלת עדיפות במהלך ההמתנה (במסגרת 20 שניות) , והגדלה שולית של העדיפות כל 60 שניות. המעקב אחר התורים מבוצע ברמות הבאות : תהליך, סוג שירות ופלח.

אופן טיפול בשיחות יוצאות :

הגדרת מושגים :

- Campaign- רכיב במערכת המכיל אוסף רשומות לפניות ללקוחות
- Campaign list- אוסף רשומות המוזרמות לכל אחד מן ה-campaign.
- Agent group- קבוצת עובדים המוגדרים לעבוד למול ה-campaign
- Acd group- קבוצת ה-ACD שמוגדרת לעבודה למול ה-campaign, אשר משמשת לביצוע החיוג.
- שיטות החיוג, הניתנים להגדרה ברמת ה-campaign

Preview o

Power o

Predictive o

הקמת ה-campaign ע"י הגדרה והזרמת הרשומות המיועדות לטיפול. (הרשומות יהיו בעלי מכנה משותף לדוגמא : סוג הזדמנות או סוג מטלה).

1. הקמה והגדרה של קבוצת הנציגים-agent group- שתטפל ב-campaign .

2. השיחות הרשומות ב-campaign יועברו לנציגי השירות המוגדרים בקבוצת הנציגים המשויכת בלבד. במסגרת הקמת קבוצת הנציגים יתקיימו שתי אופציות :

- Blending – שילוב של שיחות נכנסות יוצאות לפי מצב התורים (הגדרות שמבוצעות במערכת).

- Out bound – חיוג יוצא בלבד

שיוך הנציגים לקבוצות יתחשב באופציות אלו, כאשר מוקדן העובד במצב blending יוגדר במסגרת ה-in bound ובמסגרת קבוצת ה-campaign. ואילו מוקדן העובד ב-out bound יוגדר בקבוצת ה-campaign בלבד.

5.2 ניתוב מבוסס מיומנות בחברת פלאפון -

5.2.1 הקדמה -

פלאפון תקשורת הוקמה בשנת 1986 על ידי מוטורולה והייתה החברה הראשונה שהחזירה את השירות הסלולרי לשימוש בישראל. משרדה הראשי של החברה ממוקדמים בגבעתיים ו-24 מרכזי שירות ללקוחות פזורים ברחבי הארץ לנוחות לקוחות החברה פלאפון מעסיקה כ-4,500 עובדים, רובם מהנדסים, טכנאים ואנשי שירות לקוחות. לחברה כשני מיליון לקוחות פעילים.

לקוחות החברה נהנים משירות מקצועי ויעיל של צוות נציגי מוקד שירות הלקוחות. נציגי המוקד מוסמכים להעניק את השירות הטוב ביותר בזמן הקצר ביותר, 24 שעות ביממה. הנציגים דוברי עברית, אנגלית, ערבית ורוסית ומתודרכים לתת שירות בנושאים שונים ומגוונים.

בנוסף למוקד הטלפוני, מפעילה פלאפון גם מרכזי שירות לקוחות הפרוסים ברחבי הארץ. מרכזי השירות מספקים מגוון של שירותים תחת קרת גג אחת.

פלאפון התחייבה לשירות מהיר במוקדי ומרכזי השירות – "30:30": מחויבות למענה עד 30 שניות במוקדים הטלפוניים ולטיפול של עד 30 דקות במרכזי השירות. פלאפון מספקת ללקוחותיה שירות גם באמצעות אתר האינטרנט (שירות Online). באפשרות לקוחות החברה לקבל פרוט שיחות, מצב חשבון ומידע נוסף.

פלאפון פועלת באמצעות שני מותגים במטרה לספק לקהלי יעד שונים חווית תקשורת ייחודית על פי צרכי קהל היעד. מותג Esc שהושק ב-2003 הוא כיום המותג המוביל בסגמנט הצעירים וה-Mobile Entertainment.

5.2.2 מערך השירות בחברת פלאפון -

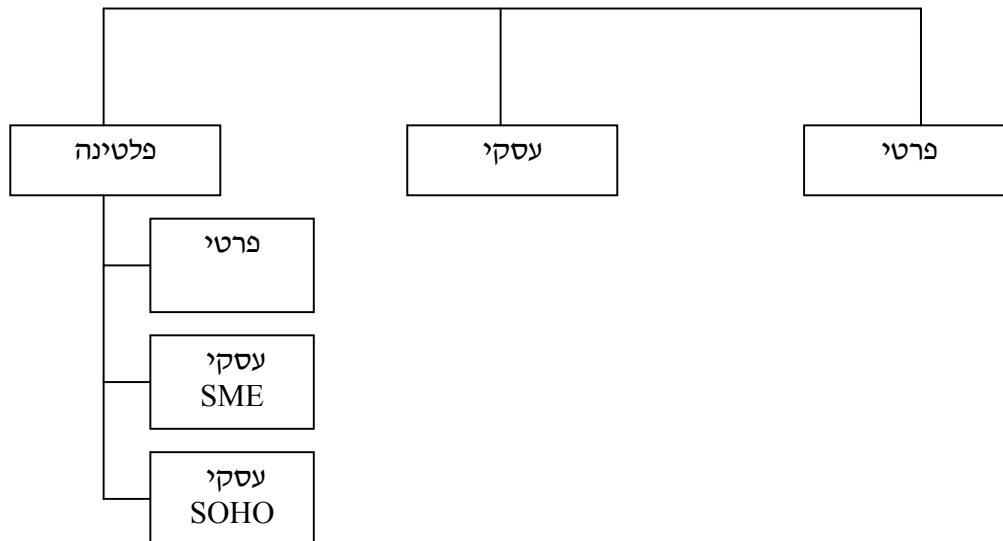
כאמור, מערך השירות של חברת פלאפון פועל בשני מישורים – מוקדי שירות טלפוניים ומרכזי שירות לקוחות. בפרויקט זה נתעסק במערך השירות הטלפוני של חברת פלאפון. קיימים שלושה מרכזי שירות טלפוניים בחברה – המוקד המרכזי יושב במשרדים החברה בגבעתיים, והמוקדים הנוספים ממוקמים ברמת גן ובבאר שבע.

פלאפון נותנת שירות ללקוחות במספר תחומים – שירותים כלליים, שירותים פיננסיים, שירותי תוכן ושירותים טכניים. על שירותים אלו יפורט בהמשך.

ההגעה למוקדי השירות היא באמצעות מספרים מקוצרים ממכשירי החברה או מקווי בזק. קיימים מספר מספרים עבור המוקדים השונים.

5.2.3 סיווג הלקוחות -

החלוקה העיקרית של לקוחות החברה נעשית על פי ערך הלקוחות. בתרשים 1 מוצגת חלוקת הלקוחות בסיווג זה:



תרשים 1

- לקוחות עסקיים – מדובר בעסקים גדולים הנמצאים בהסכם עם חברת פלאפון על מתן שירותי תקשורת לעובדי העסק במחירים מוזלים.
- לקוחות עסקי SME – עסקים קטנים אשר מקבלים שירות מחברת פלאפון.
- לקוחות עסקי SOHO – עסקים בינוניים אשר מקבלים שירות מחברת פלאפון.

ללקוחות הפלטנה עדיפות על שאר הלקוחות. את הביטוי לעדיפות זו נראה בהמשך.

5.2.4 מוקדי השירות וההתמחויות -

פלאפון, כאמור, מספקת שירות תמיכה טלפוני במספר תחומים. לפיכך הקימה החברה מספר מוקדים אשר כל אחד הוא בעל התמחות שונה:

- מוקד שירות לקוחות פרטיים – מוסמך לטפל בלקוחות מסוג פרטי בנושאים כלליים. לשרתים יש את היכולת לשרת לקוחות מסוג פרטי פלטנה.
- מוקד אמיר – מטפל בלקוחות פרטי פלטנה ולקוחות SME בכל נושא אפשרי. מוסמך גם לטפל בלקוחות SOHO.

- מוקד טל סוהו – מטפל בלקוחות SOHO בכל נושא אפשרי. יכול לטפל בלקוחות SME ולקוחות עסקיים.
- מוקד טל עסקי – מטפל בלקוחות עסקיים בנושאים כלליים בלבד, ויכול לטפל בלקוחות SOHO.
- בנוסף למוקדים המחולקים ע"פ ערך הלקוח, קיימים מספר מוקדים מקצועיים:
- מוקד הנותן שירותים בשפה הערבית בכל הנושאים. מוקד זה יכול לקבל שיחות מכל המוקדים האחרים
- מוקד הנותן שירותים בשפה הרוסית בכל הנושאים. מוקד זה יכול לקבל שיחות מכל המוקדים האחרים.
- מוקד שע"מ – טיפול בבעיות גלישה ותוכן (למשל הורדות שונות מאתר האינטרנט של החברה).
- מוקד כספים.
- מוקד איזי (מכשירי טלפון אשר מופעלים על כרטיס המוטען בסכום קבוע מראש).
- מוקד טכני ותפעול מכשירים.

5.2.5 הכשרת נציגי השירות -

נציגי השירות של חברת פלאפון מאותרים ע"י חברת כ"א. לאחר מעבר של ראיונות ומבחנים הם מגיעים לקורס בחברה בו הם לומדים את המיומנויות השונות. נציג שירות מגויס למוקד שירות מסוים וע"פ הגיוס הוא עובר את הקורס המתאים. לאחר תקופה מסוימת יש באפשרות הנציגים להתקדם לתפקיד של נציג שירות באחד ממוקדי הפלטינה. משך העבודה הממוצע של נציג שירות הוא כשנה וחודשיים, ולכן קיימת תחלופה גבוהה של נציגים במוקד, מה שעלול להוריד את רמת השירות של המוקד.

5.2.6 רמת השירות -

כאמור, פלאפון התחייבה בפני לקוחותיה על מענה טלפוני של 30 שניות. על מנת לשמור על יעד זה הגדירה לעצמה חברת פלאפון יעדי שירות. המדד לשמירה על יעדים אלו הוא שיעור הלקוחות המקבלים שירות תוך זמן היעד עליו התחייבה פלאפון – 30 שניות. מדד זה שונה בין סוגי הלקוחות השונים, ככל שערך הלקוח גבוה יותר כך גדל שיעור הלקוחות שיקבלו שירות בזמן המטרה.

5.2.7 נטישות המערכת -

קיימים שני סוגים של נטישות:

1. לקוחות אשר נוטשים בזמן מענה ה-VRU – לקוחות אלו לא נספרים על ידי הארגון כנוטשים. ההנחה המביאה לכך היא שלקוחות אלו לא נוטשים בגלל השירות אלא בגלל סיבות אישיות שלהם.
2. לקוחות הנוטשים לאחר ביום ה-VRU – לקוחות אלו נספרים במערכת כ"נוטשים" את המערכת, אך לא נעשה כל שימוש בנתון זה על מנת לקבוע את הפרמטרים של המערכת.

5.2.8 הגעת הלקוחות למוקדי השירות -

- לקוחות מגיעים למוקדי השירות ע"י חיוג מכשיר הפלאפון שברשותם או ממכשירי בזק. אשר לקוח מחייג ממכשיר של חברת פלאפון המערכת מזהה אוטומטית את המכשיר ממנו בוצעה השיחה, ואילו לקוח המתקשר לקבלת שירות מטלפון קווי מתבקש לחייג את מספר הפלאפון לו נדרש השירות.
- למוקדים פרטי, עסקי ופלטניה לסוגיו מתקשרים 166*. מערכת הניתוב מזהה את סוג הלקוח ומנתבת את השיחה למוקד המתאים.
- לקוחות פלטניה מקבלים הודעה המאפשרת להם לבחור בין מענה אוטומטי (למשל על מנת לקבל מצב חשבון או נציג שירות. כאמור, נציגי השירות במוקד הפלטניה יכולים לשרת את הלקוחות בכל נושא שירצו.
- לקוחות פרטיים או עסקיים מקבלים תפריט לניתוב שיחות ע"פ סוג השירות המבוקש. באם הם מבקשים שירות מתמחה (כספים, שע"מ וכו') המערכת מעבירה אותם אוטומטית למוקד המתאים. אחרת המערכת מעבירה אותם למוקד הכללי.
- ישנה אפשרות להתקשר היישר לאחד מהמוקדים המתמחים (שפות, כספים וכו') ע"י מספר שונה. כאשר לקוח מגיע למוקד מתמחה הוא נכנס לתור ללא עדיפות כלשהי.
- כאשר לקוח מתקשר למוקד הכללי ומבקש נציג שירות, ולאחר מכן הוא מבקש טיפול בנושא השייך לאחד מהמוקדים המתמחים, הוא מועבר לתור של המוקד המתמחה המבוקש.

5.2.9 מבנה המחלקה -

המחלקה האחראית לתפעול מוקד השירות מורכבת משני צוותים :

1. צוות מחשבים וטלפוניה, אשר בראשו עומד רוני רומנו שהנחה אותנו בנעשה בפלאפון. צוות זה אחראי על תפעול מערכת הניתוב "אספקט". בין התפקידים של הצוות ניתן למנות :

1. קביעת פרוטוקולי הניתוב כפי שיוסבר בסעיף 5.2.10. קביעת הפרוטוקולים נעשית בעיקר בשיטה של "ניסוי וטעייה". המערכת נתמכת ע"י תוכנת מחשב אשר באמצעותה אפשר לקבוע את התסריטים השונים לניתוב הלקוחות.

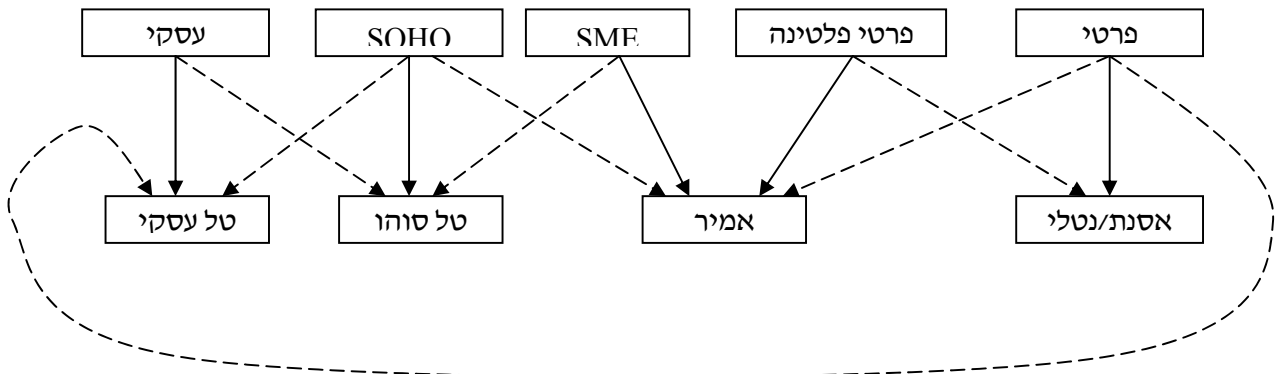
2. הגדרת מיומנויות הנציגים במערכת ע"פ דרישות שנקבעו.

3. הגדרת עדיפויות לסוגי הלקוחות השונים ע"פ פרוטוקולי הניתוב.

2. הצוות השני אחראי על התחום האנושי. הוא אחראי על דיוס והכשרת נציגי שירות על פי הדרישות למיומנויות השונות ועל פי התחזיות לדרישות, ושיבוץ כ"א במשמרות על פי האילוצים השונים. התוכנה איתה עובד צוות זה היא תוכנת ESI לניהול ושיבוץ כ"א.

5.2.10 ניתוב הלקוחות -

בתרשים 2 מוצג המודל המתאר את גלישות הלקוחות האפשריות בין המוקדים השונים.



תרשים 2

כפי שניתן לראות מהתרשים, קיימת תופעה של גלישות במערכת הניתוב. גלישות מתקיימות במקרה העומס על המוקדים השונים אינו מאוזן, תוך שימוש בתנאי סף שיוסברו בהמשך. נדגים ניתוב אופייני של לקוחות פרטיים (ללא לילה וסופי שבוע):

שלב 1 – במידה ואין פנויים, ויש לפחות שלושה שרתים פנויים במוקד אמיר, הלקוח יועבר למוקד אמיר.

שלב 2 – אם הלקוח לא נענה, המערכת בודקת במשך 4 פעמים כל 3 שניות אם (א) יש פנויים במוקד אסנת/נטלי (ב) יש לפחות שלושה שרתים פנויים במוקד אמיר.

שלב 3 – אם הלקוח לא נענה, המערכת בודקת במשך פעמיים כל 4 שניות אם יש לפחות שלושה שרתים פנויים בטל עסקי.

שלב 4 – אם הלקוח לא נענה, הוא ממתיך 16 שניות בתור של מוקד אסנת/נטלי, אלא אם יש לפחות שני שרתים פנויים במוקד אמיר.

שלב 5 – אם הלקוח לא נענה, הוא ממתיך 25 שניות בתורים של המוקדים אסנת/נטלי, אמיר ואיזי.
שלב 6 – אם הלקוח לא נענה, הוא מקבל עדיפות 8 (*) וממתין באותם התורים שבשלב 5 למשך 60 שניות.

שלב 7 – אם הלקוח לא נענה, הוא ממתיך בכל התורים, ולאחר 30 שניות המתנה הוא מקבל עדיפות 7.

(*) לקוח פרטי המגיע למערכת מקבל אוטומטית עדיפות 9 שהיא העדיפות הנמוכה יותר. כאשר הלקוח ממתיך זמן רב הוא מקבל עדיפויות גבוהות יותר (8 ולאחר מכן 7) על מנת שיקבל את השירות במהרה.

דוגמא נוספת היא הגלישה של לקוח פלטינה למוקד הפרטי. במקרה והלקוח ממתין בתור למוקד אמיר למעלה מ-28 שניות הוא יעבור לתור השייך למוקד אסנת / נטלי תוך קבלת עדיפות 8. העדיפות תגרום ללקוח הפלטינה לעבור "לעקוף" את כל הלקוחות הפרטיים ובכך נקטין את זמן ההמתנה שלו.

עבור כל סוג לקוח, קיימים ניתובים שונים בין המוקדים השונים. לא מציג את הניתובים המלאים עקב בקשת "פלאפון". כמו כן קיימים תסריטי ניתוב שונים עבור זמנים שונים – קיימת הבדלה בין שתי קבוצות זמנים:

1. ימים א'-ה' בשעות 07:00-24:00, ויום ו' בשעות 07:00-15:00 (ימים רגילים)
 2. ימים א'-ה' בשעות 07:00-24:00, והחל מיום ו' בשעה 15:00 עד יום א' בשעה 07:00 (לילות וסופ"ש). בסעיף זה נכנסים גם חגי ישראל במתכונת של סופי שבוע.
- הניתוב בין קבוצות זה שונה שכן העומס בקבוצה 1 גבוהה בהרבה מזה שבקבוצה 2. בזמנים השייכים לקבוצה 2 מוקד נהפך ל"סופר-גרופ" בו כל השרתים נותנים שירות לכל הלקוחות.

פרק 6 –

– ניתוב מבוסס מיומנות

ניתוח מודלים

6.1 מודל V -

6.1.1 הגדרת המודל והתסריט הנבדק -

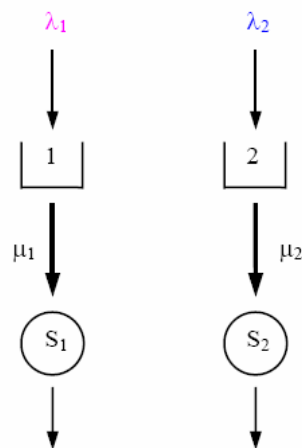
כאמור, ניתוב מבוסס מיומנויות מתייחס למערכת בה לקוחות שונים דורשים סוג שירות שונה, ולמוקדנים יש התמחויות שונות. ניתוב מבוסס מיומנויות מחייב שלוש החלטות מרכזיות:

1. חלוקת הלקוחות לסוגים שונים.
 2. חלוקת מוקדנים ע"פ מיומנויות.
 3. ניתוב מבוסס מיומנויות של לקוחות מסוגים שונים לשרתים השונים.
- סוגי לקוחות יכולים לייצג למשל לקוחות הדורשים שירות שונה כמו תמיכה טכנית או הנהלת חשבונות, וכן לקוחות בעלי חשיבות שונה מנקודת המבט של הארגון למשל לקוחות VIP לעומת לקוחות רגילים. לעומת זאת סוגי שרתים שונים יכולים לייצג למשל שרתים דוברי שפות שונות או מומחים לעומת טירונים וכן שרתים בעלי הכשרה כללית לעומת כאלה בעלי הכשרה ספציפית.
- ניתוב מבוסס מיומנויות הוא פרוטוקול לניתוב בזמן אמת וניתן לאפיון בשתי דרכים:

1. ניתוב דינאמי- הניתוב נקבע על פי מצב המערכת בכל נקודת זמן.
 1. ניתוב סטאטי- הניתוב נקבע על פי כללים קבועים מראש שאינם משתנים בהתאם למצב המערכת.
- מודל ה-V בו נדון הוא סוג של מודל ניתוב לקוחות דינאמי.

כפי שראינו בפרק 4.2, המאמר מציג את מודל I^2 שמתאר שתי מערכות מקבילות, אשר את כל אחת

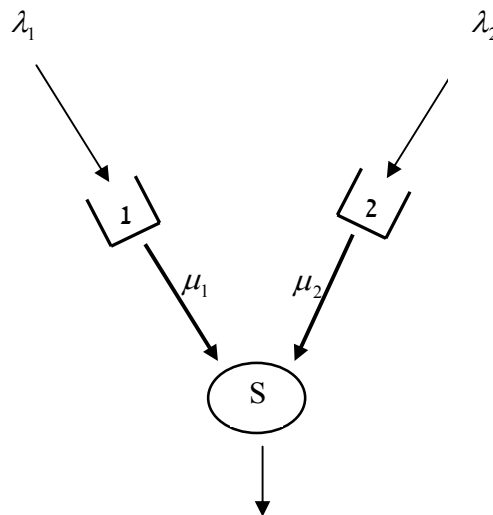
מהן ניתן לתאר כמערכת M/M/N. מודל I^2 מתואר בתרשים 1.



תרשים 1

נניח כעת שלקבוצת השרתים s_2 מגיעים לקוחות רגילים ואילו לקבוצת השרתים s_1 מגיעים לקוחות VIP, כאמור לקוחות שהערך הכספי שלהם מנקודת המבט של הארגון הוא גבוה יותר. במודל ה-V אנו דנים באפשרות לאחד את שתי קבוצות השרתים המופיעים במודל ה- I^2 לקבוצה אחת המכילה שרתים המסוגלים לטפל בלקוחות מכל סוג, אולם קיימת מדיניות ברורה הקובעת באיזה סוג לקוח יש לטפל בכל נקודת זמן על פי מצב המערכת.

המבנה של מודל ה-V מתואר בתרשים 2.



תרשים 2

הנחת המודל היא שלקוחות ה-VIP אמורים לקבל שירות טוב יותר, כיוון שהם שווים יותר. לכן נציע מדיניות סף לפיה ניתנת עדיפות ברורה ללקוח VIP לעומת לקוח רגיל:

- בכל עת השרתים ב-S ישרתו את לקוחות ה-VIP אם קיימים כאלה במערכת.
- במידה ומספר השרתים שמתפנה עולה מעל סף מסוים, הלקוחות הרגילים יקבלו שירות.

מכאן ברור שכמות הזמן שלקוח רגיל יאלץ להמתין בתור תהיה גדולה יותר בממוצע לעומת לקוח VIP. אנו נבחן כאן את מידת ההשפעה של המדיניות הזו על זמני ההמתנה של הלקוחות כולם ושל כל סוג לקוח בנפרד בפרט. נדגיש כאן שניתוח המודל כפי שאנו מציגים אותו מביא בחשבון את זמני ההמתנה של הלקוחות, ההשפעה של עלויות סוגי הלקוחות השונים (כתוצאה מזמני המתנה), היחסים

שבין עלויות אלה, והתועלת המתקבלת לארגון בכל מצב. הניתוח בוצע באמצעות מודל סימולציה שנכתב בשפת ++C עבור המודל וגיליון excel אליו נפלטו הנתונים.

נשתמש באסטרטגית שירות QED (Quality and Efficiency Driven) המנסה ליצור רמה גבוהה של שירות הניתן ללקוחות ובמקביל רמה גבוהה של יעילות המערכת. כלל זה יעזור לנו בקביעת היחס שבין קצב הגעת הלקוחות ובין רמת האיוש של המערכת. תוצאות יראו שמודל הסף אכן מביא למינימיזציה של העלות הכללית המתחשבת בזמני ההמתנה של שני סוגי הלקוחות ומכאן, בעלות זמני המתנה של הלקוחות. המטרה היא מציאת קשר ברור בין רמת האיוש של השרתים (מספר שרתים במערכת) לבין הסף הנבחר ע"פ האסטרטגיה שתוארה.

אנו מניחים כעת, כי לזמן ההמתנה של לקוח יש עלות, וזו מוגדרת עבור כל יחידת זמן שהלקוח ממתיין לשירות. עלויות אלו נקבעות על פי הערך של הלקוחות מהסוגים השונים לחברה. ברור שעלות זמן ההמתנה של לקוח VIP גבוהה יותר מעלות זמן ההמתנה של לקוח רגיל. נגדיר אם כן את העלויות:

C_{vip} - עלות ההמתנה ליחידת זמן של לקוח מסוג VIP.

C_{reg} - עלות ההמתנה ליחידת זמן של לקוח רגיל.

עבור יחס מסוים בין שתי העלויות האלה $\frac{C_{vip}}{C_{reg}}$ אנו מחפשים את אותו סף שיביא את העלות הכללית ממוצעת למינימום. עלות זו ניתן להגדיר באופן הבא:

$$Cost = E(wait_{vip}) \cdot C_{vip} \cdot \lambda_{vip} + E(wait_{reg}) \cdot C_{reg} \cdot \lambda_{reg}$$

כפי שניתן לראות מההגדרה זוהי בעצם העלות הממוצעת ליחידת זמן הנובעת מהמתנה של לקוחות. ההשערה אותה אנו באים לבדוק היא האם אכן קיים קשר ברור בין מספר השרתים במערכת (N) ובין הסף הנבחר. על סמך מחקרים תיאורטיים קודמים, אנו מניחים בשלב זה כי הסף האופטימלי יהיה מסדר גודל של $\log(N)$.

את קצב הגעת הלקוחות אנו קובעים על פי כלל השורש לאיוש (Helffin QED Teorem-Witt 1981):
 $N = R + \beta \cdot \sqrt{R}$
 המופיע בתיאורית ה QED (תיאור מקוצר של התיאוריה מופיע בנספח 2)
 המאפשרת ליצור שילוב של רמת שירות ויעילות גבוהה של המערכת. על פי התיאוריה ידוע שבגבול אמור להתקיים:

$$\lim_{N \rightarrow \infty} \sqrt{N}(1 - \rho_N) = \beta$$

$$0 \leq \beta \leq \infty$$

לשם פשטות, נקבע את קצב הגעת הלקוחות μ להיות 1 מכאן נקבל שמתקיים $\rho_N = \frac{\lambda}{N}$ ולכן

בקירוב נוכל לומר שצריך להתקיים: $\lambda = N - \beta \cdot \sqrt{N}$. על פי הנוסחה הנ"ל ניתן לקבוע את קצב

הגעת הלקוחות למערכת עבור מספר נתון (N) של שרתים וציון מסוים של רמת שירות, אותו נבחר לשם פשטות להיות $\beta = 1$.

את בדיקת ההשערה שלנו אנו מבצעים באמצעות סימולציה עבור שלושה תסריטים שונים של קצבי הגעה. ניתוחי עלות תועלת יבוצעו בהמשך עבור התסריט הקרוב ביותר להתנהגות המערכת במציאות מבין שלושת המקרים שנבחנו. אנו מבחינים בין המקרים הבאים:

- קצב ההגעה של לקוחות VIP שווה לקצב ההגעה של לקוחות רגילים -

$$\lambda_{vip} = \lambda_{reg} = \frac{1}{2} \lambda$$

- קצב ההגעה של לקוחות VIP גדול פי 2 מקצב ההגעה של לקוחות רגילים -

$$\lambda_{vip} = 2 \cdot \lambda_{reg} \Rightarrow \lambda_{vip} = \frac{2}{3} \lambda, \lambda_{reg} = \frac{1}{3} \lambda$$

- קצב ההגעה של לקוחות VIP קטן פי 2 מקצב ההגעה של לקוחות רגילים -

$$\lambda_{vip} = 0.5 \cdot \lambda_{reg} \Rightarrow \lambda_{vip} = \frac{1}{3} \lambda, \lambda_{reg} = \frac{2}{3} \lambda$$

6.1.2 הגדרת הסימולציה -

הנחות הסימולציה -

- משטר התור בתוך קבוצת העדיפות הוא FCFS.
 - כל הזמנים, הן זמני השירות והן הזמנים הבין מופעיים הם משתנים מקריים אקספוננציאליים בלתי תלויים.
 - בסימולציה אין התייחסות למצב של נטישות או הגעות חוזרות למערכת.
 - בסוף כל סימולציה חושבו ממוצעי זמני ההמתנה בתור עבור כל אחד מסוגי הלקוחות. מאחר וכל הזמנים שהוגרלו נלקחו מההתפלגות האקספוננציאלית, השתמשנו בתכונת חוסר הזיכרון של ההתפלגות, ובכל איטרציה הוגרלו ערכי הזמן של שלושה אירועים:
1. הזמן עד הגעת הלקוח VIP הבא.
 2. הזמן עד הגעת הלקוח הרגיל הבא.
 3. הזמן עד שהשרת הבא מתפנה.
- "האירוע הבא" מוגדר להיות המינימום מבין שלושת האירועים הללו, והזמן המתאים מתווסף לזמן הנוכחי המוגדר בסימולציה ומתקדם לאורך הריצה.

כאשר הזמן הנוכחי הזה מגיע ל- 100 יחידות זמן, מה שמעיד על כך שהסימולציה מתקרבת למצב יציב, חישבנו בכל איטרציה את זמן ההמתנה הממוצע של כל סוג לקוח. כאמור, נניח שכעת האירוע הבא יתרחש בעוד X יחידות זמן מהזמן הנוכחי המוגדר, אזי זמן ההמתנה בכל תור שווה לזמן ההמתנה הקודם שחושב ועוד מספר הלקוחות שכעת בתור המסוים מוכפל ב- X .

מטרת הסימולציה –

המטרה היא לאשש את ההשערה כי היחס בין הסף האופטימלי לבין גודל המערכת (מספר השרתים) הוא מסדר של $\log(N)$ כאשר N – מספר השרתים. בנוסף נרצה ללמוד מתוצאות הסימולציה על ההתנהגות של הסף ביחס ליחסים השונים בין עלויות ההמתנה של לקוחות ה-VIP והלקוחות הרגילים.

מבנה הסימולציה –

- הסימולציה נערכה ע"י תכנות בשפת C++ - קוד הסימולציה מובא בנספח 3.
- נערכו שלוש סימולציות עבור כל אחד מהמקרים שתוארו לעיל. כאשר N - מספר השרתים במערכת. עבור כל תסריט שתואר נערכו הרצות עבור $N = 100, 200, \dots, 1000$.
- הסימולציה הכילה חמש רפליקציות שכל אחת מהן ארכה 100,000 יחידות זמן, כך שהמערכת הגיע למצב יציב.
- הסימולציה מבחינה בין המקרים הבאים:
 1. האירוע הבא הוא כאשר שרת מסיים את הטיפול בלקוח- במקרה זה השרת מוגדר כשרת פנוי כך שמספר השרתים הפנויים מוגדל באחד. אם קיימים לקוחות בתור ה-VIP השרת שהתפנה ישרת את הלקוח הבא בתור הזה, ומספר השרתים הפנויים מורד באחד. אחרת במידה והתור של הלקוחות הרגילים אינו ריק, וכעת מספר השרתים הפנויים גדול מסף מסוים, השרת ישרת את הלקוח הבא מתור הרגילים, ומספר השרתים הפנויים קטן באחד.
 2. האירוע הבא הוא הגעת לקוח- VIP – אם קיים כעת שרת פנוי כלשהו, השרת ישרת את הלקוח ומספר השרתים הפנויים מוקטן באחד. אם אין שרתים פנויים הלקוח מצטרף לתור המתאים.
 3. האירוע הבא הוא הגעת לקוח רגיל- אם מספר השרתים הפנויים גדול מערך הסף אחד השרתים (כלשהו) ישרת את הלקוח אחרת הלקוח מצטרף אל התור.
- כשזמן הריצה תם אנו מחשבים את זמן ההמתנה הממוצע של כל אחד מסוגי הלקוחות, באמצעות חילוק זמן ההמתנה הכללי בתור במספר הלקוחות שהמתינו בתור.

נכונות הסימולציה –

עבור המקרה בו ישנו שרת אחד בלבד קיימת נוסחאות אסימפטוטיות אנליטיות (מנדלבאום וגורביץ') אשר באמצעותן ניתן לחשב את זמני ההמתנה של הלקוחות השונים. על מנת לבדוק את נכונות האלגוריתם השוונו את התוצאות שנתקבלו בסימולציה לתוצאות התיאורטיות קיבלנו כי תוצאות אלו אכן זהות בדיוק של מליונית.

דיוק -

כאמור כדי להביא לדיוק של האומדן ערכנו חמש רפלקציות.

הפלט לסימולציה:

הסימולציה מעבירה את הנתונים הבאים ישירות לתוך קובץ excel:

1. הסף הרלוונטי המסומן באות k.

2. זמן ההמתנה הממוצע של לקוחות ה-VIP.

3. זמן ההמתנה הממוצע של הלקוחות הרגילים.

k	AvgWaitTimeV	AvgWaitTimeR
---	--------------	--------------

כדי לתת לכל סוג של לקוח עלות שונה, כאמור כדי להביא את חשיבות לקוחות ה-VIP לעומת הלקוחות הרגילים לידי ביטוי נגדיר את עלות לקוחות ה-VIP בכפולות של עלות הלקוחות הרגילים, במונחים של מספר השרתים במערכת. הבחנו בין ארבע מקרים:

$$1. C_{vip} = 2 \cdot C_{reg}$$

$$2. C_{vip} = \sqrt{N} \cdot C_{reg}$$

$$3. C_{vip} = N \cdot C_{reg}$$

$$4. C_{vip} = N^2 \cdot C_{reg}$$

תיאור העלויות במונחים של מספר השרתים במערכת מאפשר, באמצעות בחינת הסף המתקבל כאופטימאלי, להעריך עד כמה טובה יותר רמת השירות הניתנת ללקוח ה-VIP ומהי מידת ההשפעה על רמת השירות הניתנת ללקוח הרגיל.

באופן דומה, מספר השרתים מהווה פרמטר תפעולי של גודל המערכת. לכן שימוש בפרמטר הזה מאפשר לתאר את מידת העדיפות הניתנת ללקוח חשוב כפונקציה של גודל המערכת בה אנו דנים. אנו מסיקים שתיאור זה יכול להוות כלי חשוב בידי אנשי המעשה כדוגמת מנהלים מהנדסים או כל אדם האחראי על מערכת הדומה לזו המתוארת כאן. מבחינה פרקטית, מנהל האחראי על מוקד שירות טלפוני הדומה לזה המתואר כאן, יכול בקלות להגדיר לעצמו, על ידי קביעת הסף המתאים, את מידת העדיפות שיש לתת

ללקוחות החשובים יותר כפונקציה של גודל המערכת עליה הוא אחראי- מספר השרתים העומדים לרשותו.

באמצעות גיליון ה- excel חושבה העלות ליחידת זמן (Cost) עבור כל אחד מהיחסים שתוארו ועבור מספר שרתים - 100, 200, ..., 1000. $N = 100$. בכל שילוב נמצא הסף (k) האופטימאלי, כאמור זה שהביא את העלות שהוגדרה למינימום בכל אחת מ- 10 הרצות הסימולציה.

מהנתונים שהתקבלו ערכנו גרפים של הסף האופטימאלי שנמצא, כפונקציה של מספר השרתים (k) כפונקציה של (N).

6.1.3 תוצאות הסימולציה וניתוחן -

כפי שצינו כל הנתונים הועברו ל- excel ובוצעו החישובים שהוסברו בסעיף הקודם. לאחר מכן בנינו טבלאות המסכמות את הסף האופטימלי עבור כל N. בטבלה 1 מוצגים הנתונים האמורים עבור המקרה

$$\text{הראשון } \lambda_{vip} = \lambda_{reg} = \frac{1}{2} \lambda :$$

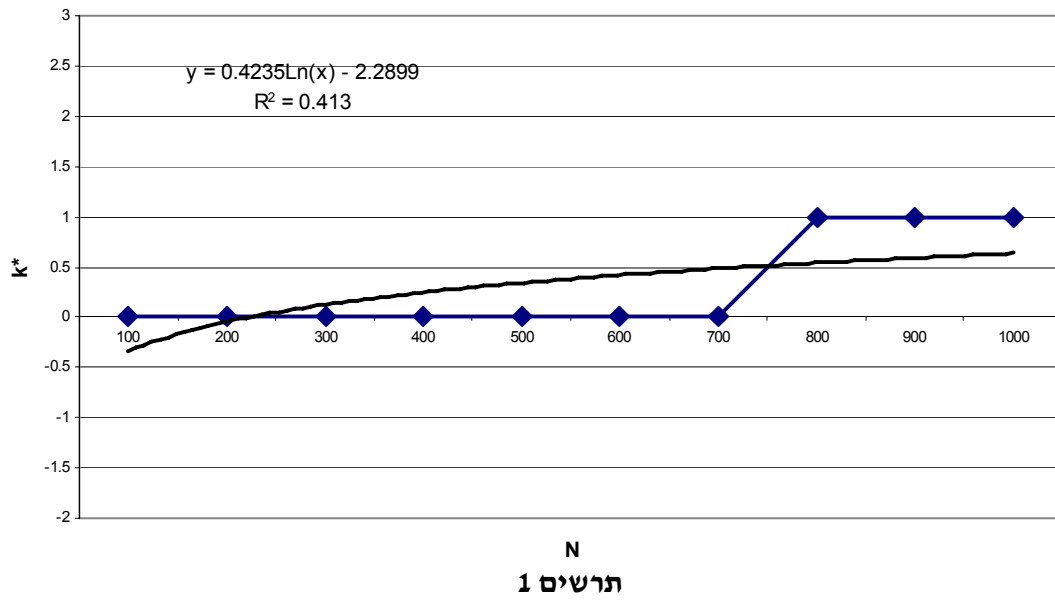
	Minimal k			
	$C_{vip}=2*C_{reg}$	$C_{vip}=\sqrt{N}*C_{reg}$	$C_{vip}=(N)*C_{reg}$	$C_{vip}=(N^2)*C_{reg}$
100	0	1	3	8
200	0	2	5	10
300	0	2	5	11
400	0	2	6	13
500	0	2	7	12
600	0	3	6	14
700	0	3	6	16
800	1	4	6	16
900	1	3	6	16
1000	1	3	7	16

טבלה-1

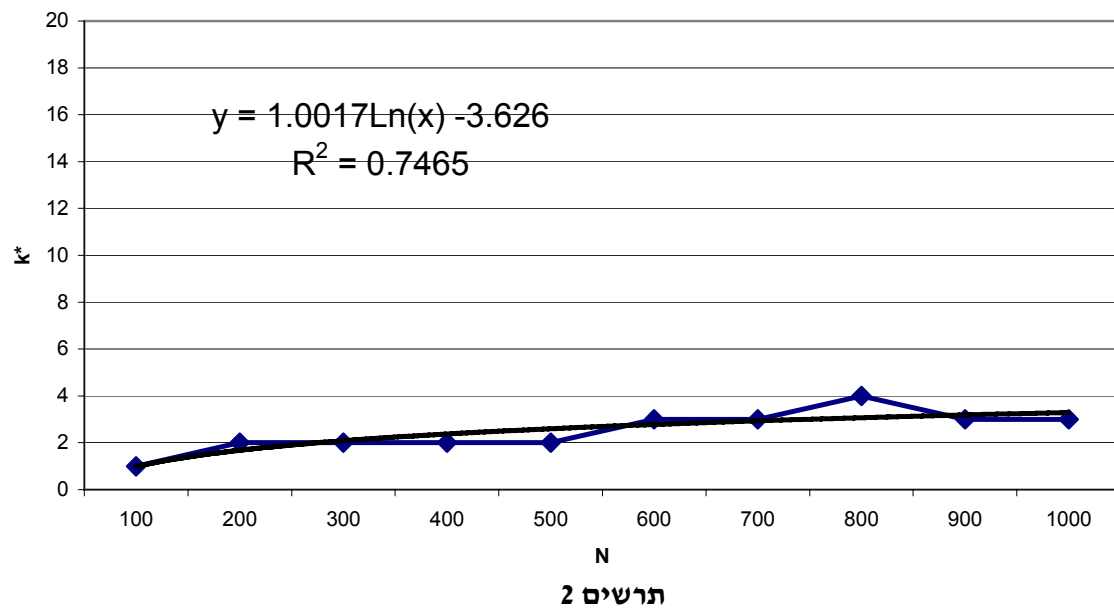
הטבלה מציגה עבור כל יחס עלות ועבור כל גודל מערכת את הסף (k) המביא לעלות מינימלית.

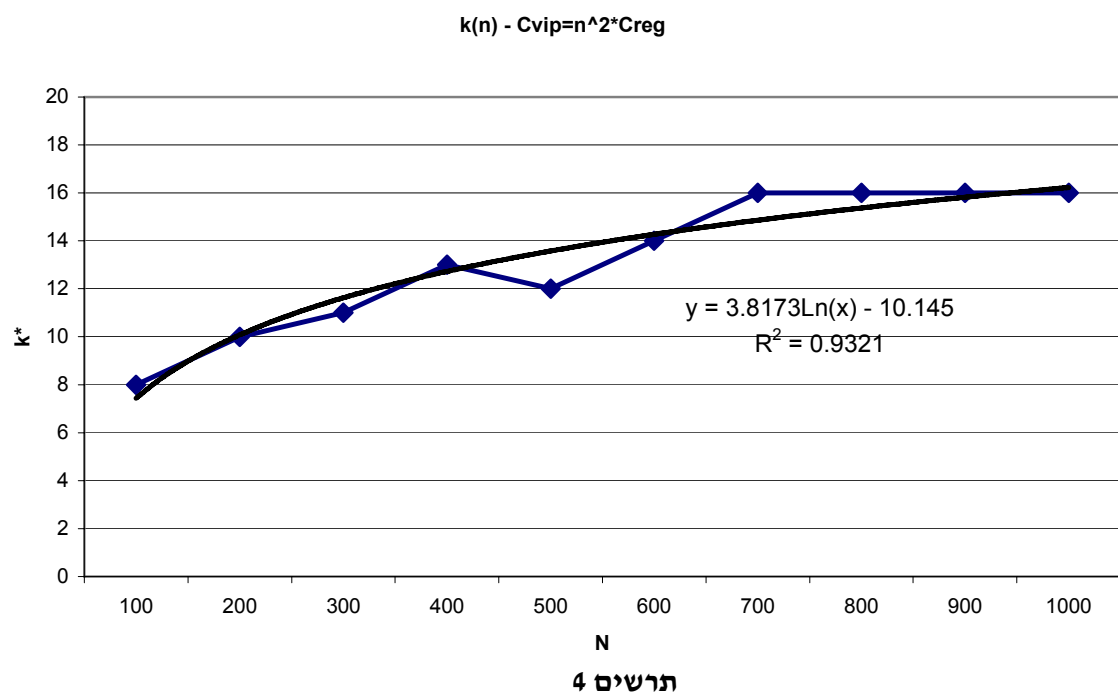
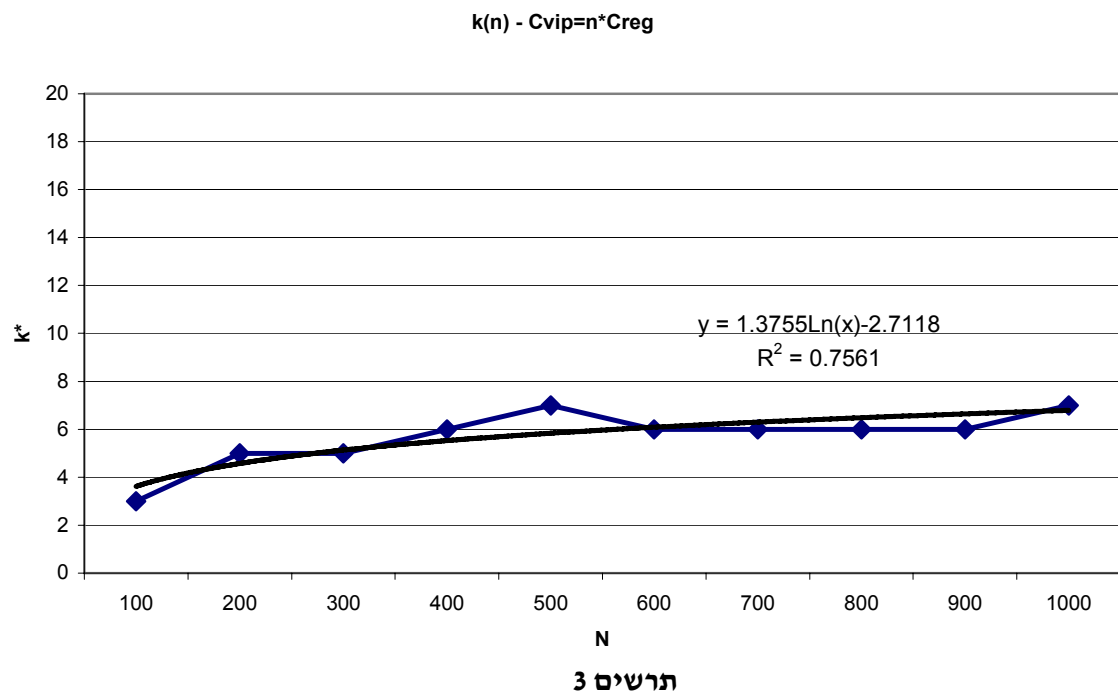
מהטבלה קיבלנו את העקומות המוצגות בתרשימים 4-1. (כל תרשים מייצג יחס שונה בין עלות ההמתנה של לקוחות ה-VIP לבין עלות ההמתנה של הלקוחות הרגילים כפי שהוסבר קודם)

$k(n) - C_{vip}=2*C_{reg}$



$k(n) - C_{vip}=\sqrt{n}*C_{reg}$

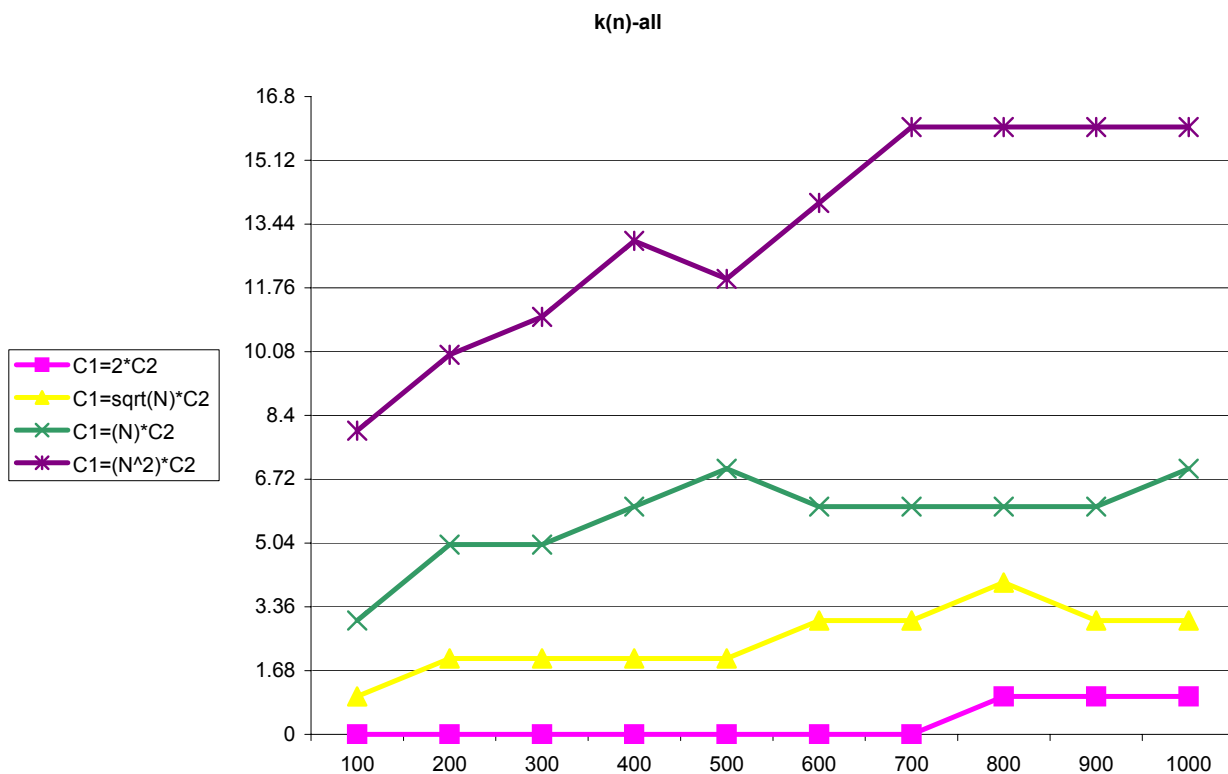




עבור כל עקומה הוספנו קו מגמה המציג את מידת הקירבה שבין הערכים שהתקבלו לבין ההשערה ששיעורנו.

ניתן לראות, עבור קצבי הגעה שווים, כי השערות מתקיימות באופן מובהק יותר, ככל שהפער בחשיבות בין לקוחות ה-VIP ללקוחות הרגילים גדל, ואכן הסף האופטימלי מבוטה בצורה טובה במונחים של $\ln(N)$.

אם נצרף את כל העקומות יחד למערכת צירים אחת, כפי שמופיע בתרשים 5, נוכל להבחין בעובדה בולטת נוספת.

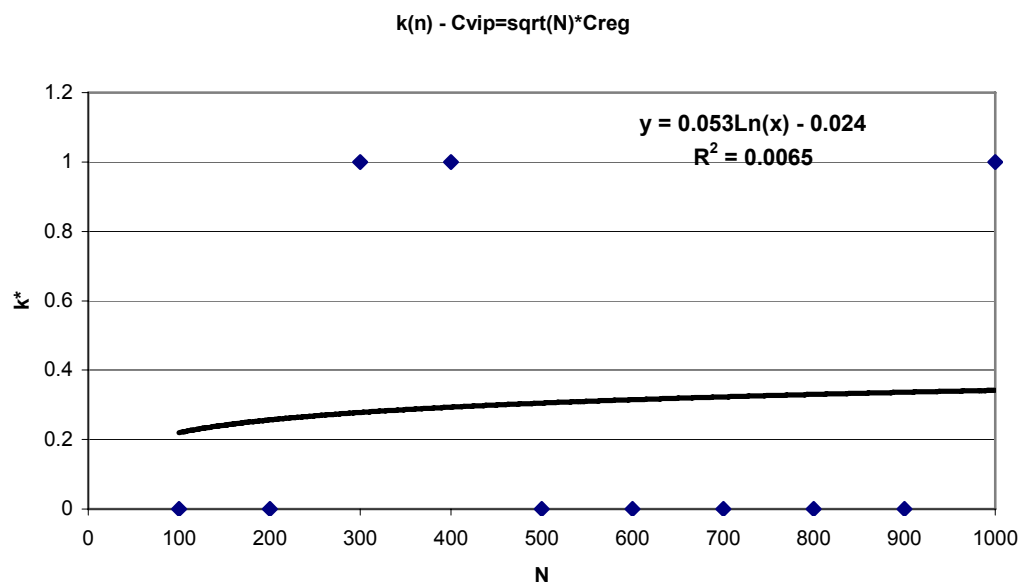


תרשים 5

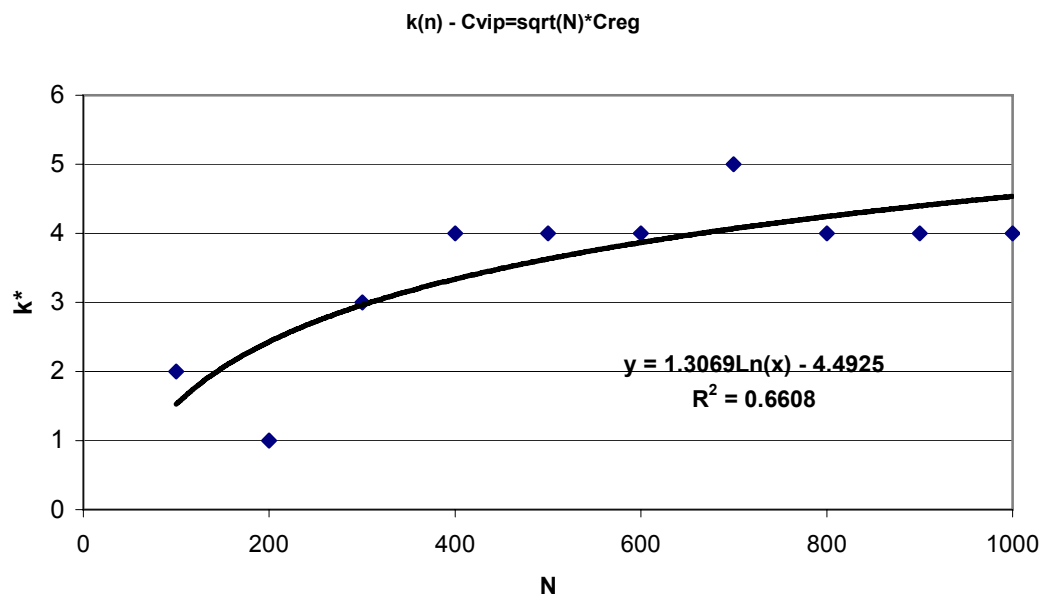
עבור הגרף התחתון ביותר, בו ההבדל בעלויות זמני ההמתנה של הלקוחות הוא נמוך יחסית (פי שתיים), מתקבל כי הסף כמעט ואינו עולה מעל לערך של אפס, ולכן במקרה זה, עבור רוב גדלי המערכת שנבדקו, מתקבל שהסף אינו תורם. לעומת זאת ניתן לראות שככל שהעקומה ממוקמת גבוה יותר, כלומר - ככל שעדיפות לקוחות ה-VIP גדלה, חשיבות הסף גדלה גם היא שכן ערכיו עולים.

בתרשימים 6-9 מוצגות התוצאות עבור המקרה השני - $\lambda_{vip} = 2 \cdot \lambda_{reg} \Rightarrow \lambda_{vip} = \frac{2}{3} \lambda, \lambda_{reg} = \frac{1}{3} \lambda$

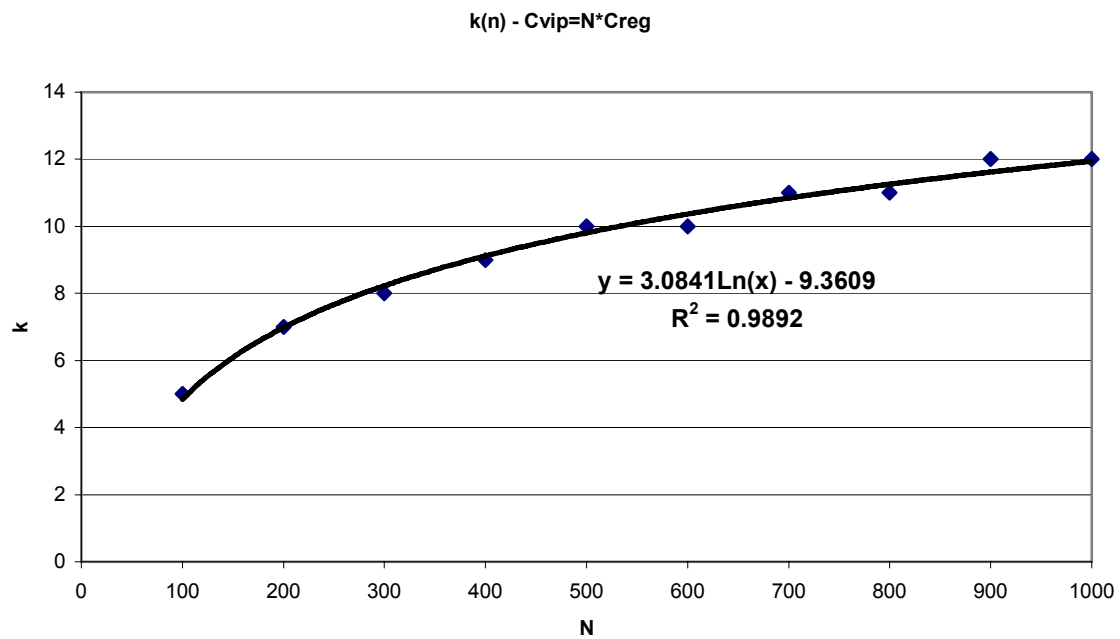
כלומר קצב ההגעה של לקוחות VIP גדול פי 2 מקצב ההגעה של לקוחות רגילים (שוב, כל תרשים מייצג יחס שונה בין עלות ההמתנה של לקוחות ה-VIP לבין עלות ההמתנה של הלקוחות הרגילים):



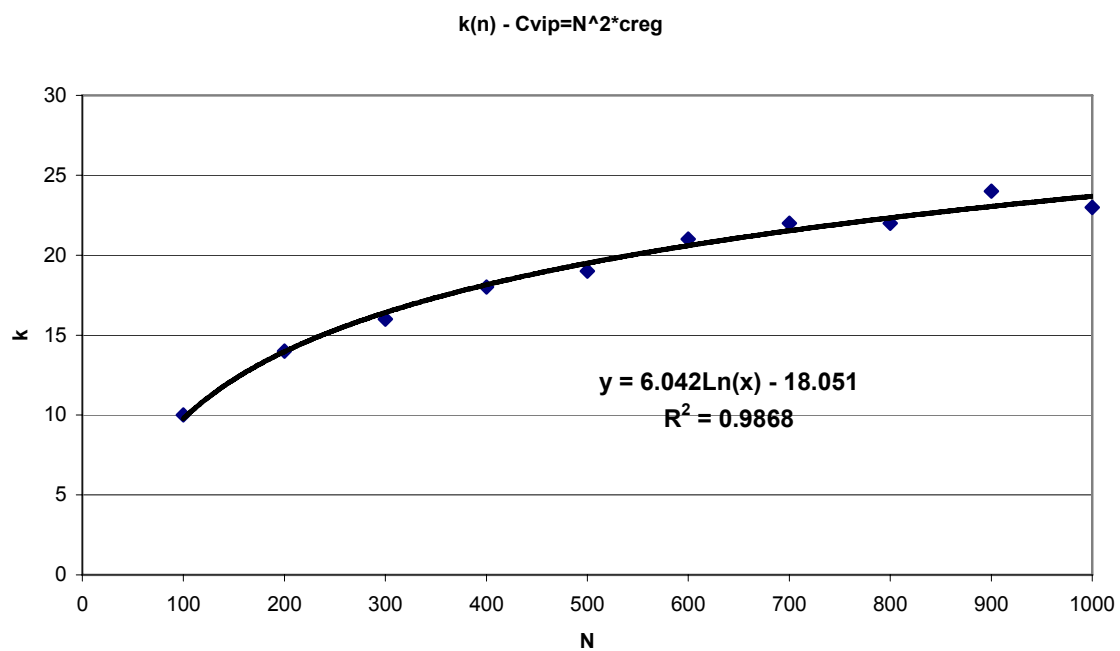
תרשים 6



תרשים 7

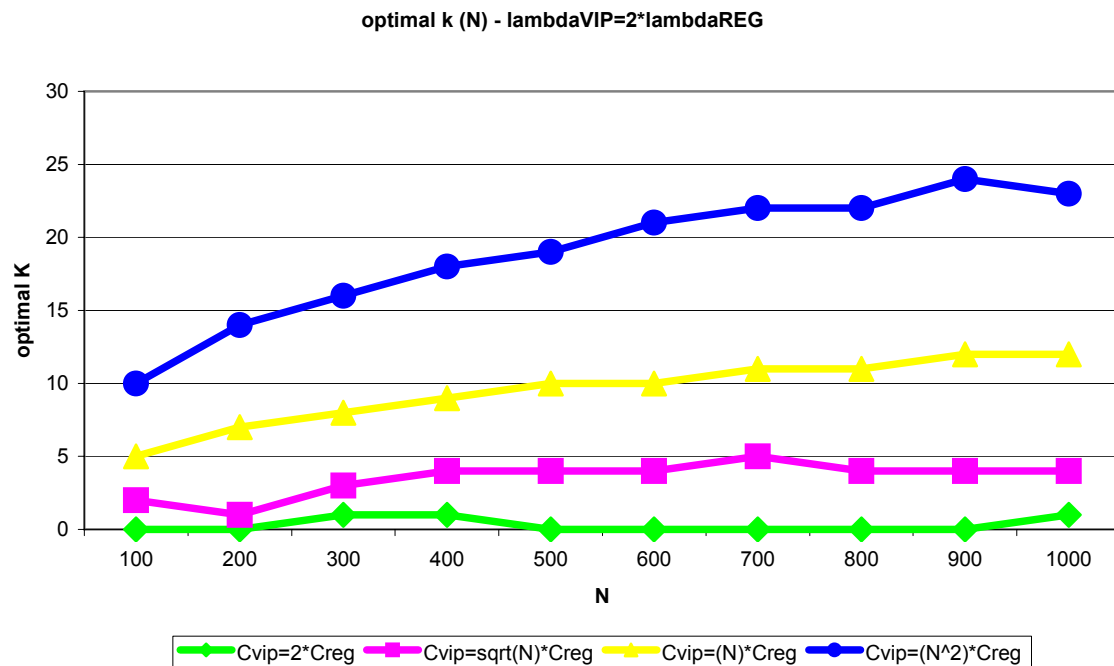


תרשים 8



תרשים 9

בתרשים 10 מופיעים הגרפים עבור כל יחסי העלויות:



תרשים 10

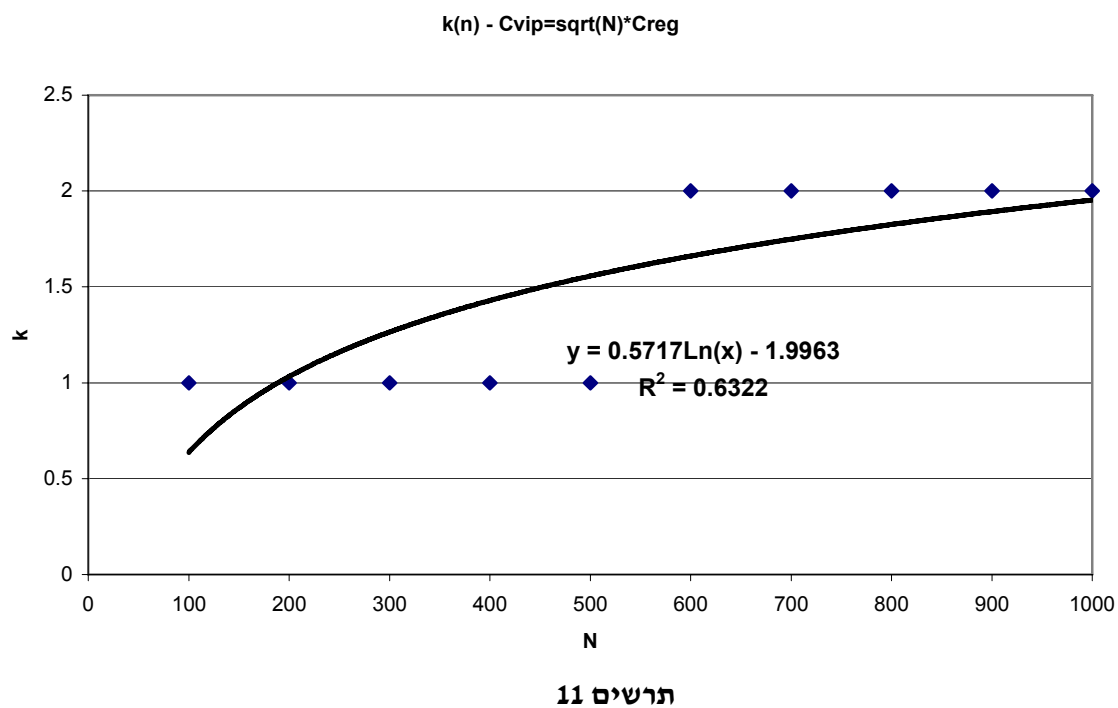
גם במקרה זה ניתן לראות שכל התופעות שתוארו במקרה הקודם, מופיעות שוב ואף באופן בולט יותר. כאמור, כאשר ההבדל בעלויות זמני ההמתנה של הלקוחות הוא נמוך יחסית (הגרף התחתון בתרשים 10), הסף כמעט ואינו עולה מעל לערך של אפס. על פי קווי המגמה בתרשימים 6-9, ההשערה ששיעורו קודם אכן מתרחשת באופן מובהק.

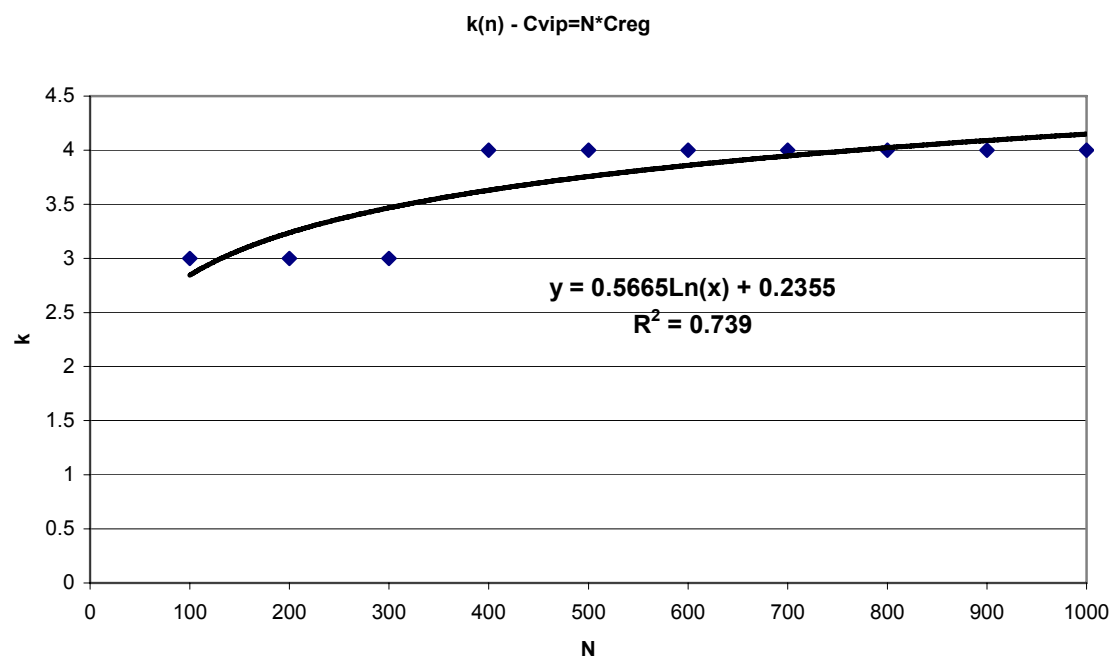
בתרשים 10 ניתן לראות שככל שהעקומה גבוהה יותר, בלומר ככל שעדיפות לקוחות ה- VIP גדלה, חשיבות הסף גם היא גדלה וערכיו עולים.

חשוב להדגיש, כי מקרה זה הוא המקרה הקרוב ביותר למציאות שכן נצפה שמספר הלקוחות הרגילים המגיעים למערכת גדול יותר ממספר לקוחות ה- VIP המגיעים למערכת.

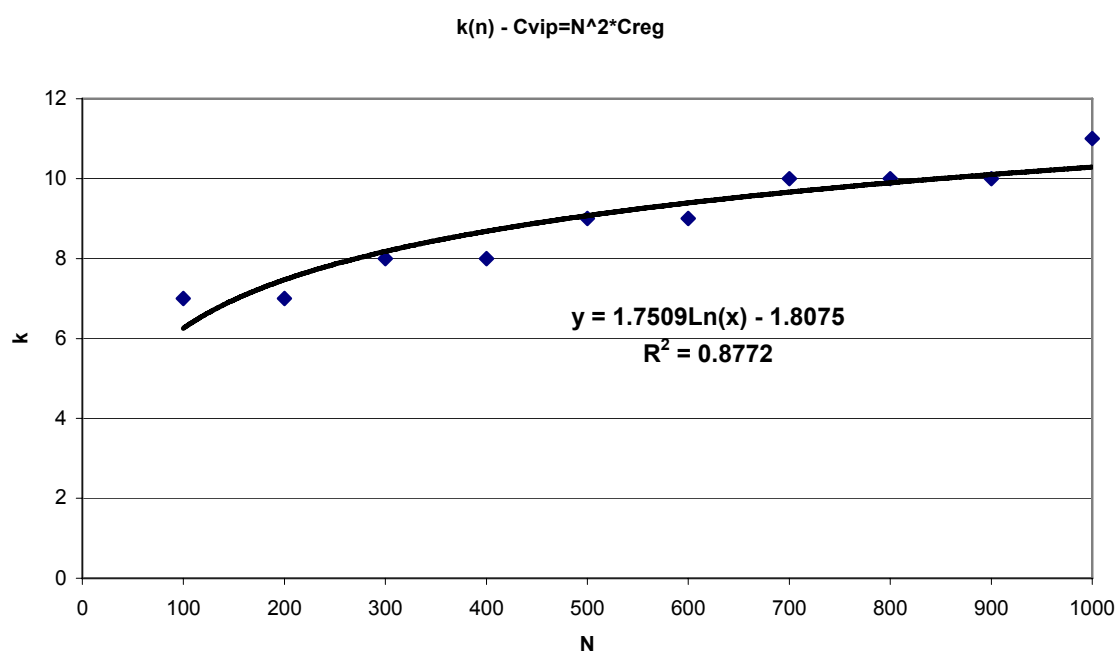
בתרשימים 11-14 מוצגות התוצאות עבור המקרה בו $\lambda_{vip} = 0.5 \cdot \lambda_{reg} \Rightarrow \lambda_{vip} = \frac{1}{3} \lambda, \lambda_{reg} = \frac{2}{3} \lambda$

כלומר, קצב ההגעה של לקוחות VIP קטן פי 2 מקצב ההגעה של לקוחות רגילים (בחלק זה לא הוצג הגרף עבור יחס עלויות 2, שכן כפיש ניתן לראות בתרשים 14 כל הערכים האופטימליים שהתקבלו שווים לאפס)

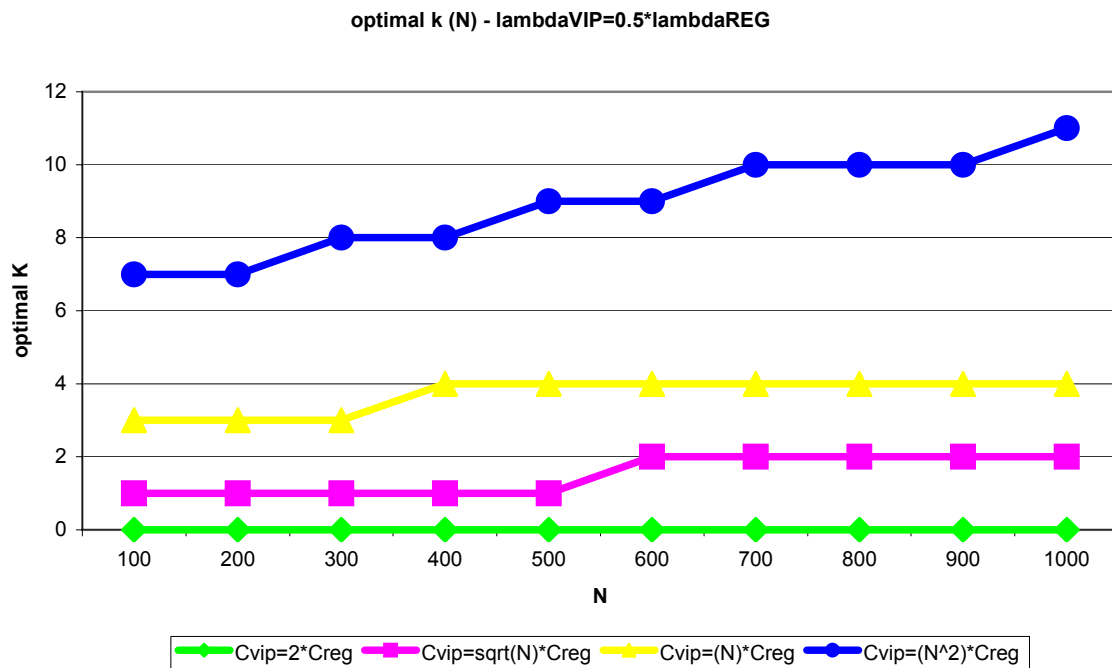




תרשים 12



תרשים 13



תרשים 14

גם במקרה זה קווי המגמה מאשרים את נכונות הטענה, אולם במקרה זה בולטת כאן העובדה שנכונות ההשערה תלויה מאוד במידת ההבדל שבחשיבות לקוחות ה- VIP לעומת הלקוחות הרגילים. כאמור, ככל שההבדל בעלויות עולה כך ה- R^2 מקבל ערך גבוהה יותר ונכונות ההשערה מבוטאת. בנוסף ניתן לראות שהסף חסר משמעות לחלוטין כשמדובר בפער קטן בחשיבות הלקוחות. נביא דוגמה פשוטה של חישובי עלות תועלת המדגישים עניין זה, בצירוף מסקנות כלליות יותר.

ניתוחי עלויות:

טבלאות 2-4 מציגות דוגמא לחישוב פשוט של ניתוח עלויות עבור היחסים N , $\sqrt{N}$ ו- N^2 . התייחסו למקרה בו קצה הגעת הקוחות הרגילים גדול פי שניים מקצב הגעת לקוחות ה-VIP, שכן כפי שהסברנו זהו המקרה המציאותי ביותר.

cost analysis (per time unit)

server	sqrt(N)					
	cost k=0	cost k*	diff	diff for day	diff per year	diff ratio
100	2.8105	2.6477	0.1628	234.432	85567.68	5.79%
200	4.1936	3.7707	0.4229	608.976	222276.24	10.08%
300	5.1813	4.6124	0.5689	819.216	299013.84	10.98%
400	6.126	5.2582	0.8678	1249.632	456115.68	14.17%
500	6.928	5.894	1.034	1488.96	543470.4	14.92%
600	7.697	6.4733	1.2237	1762.128	643176.72	15.90%
700	8.2436	6.9701	1.2735	1833.84	669351.6	15.45%
800	8.9408	7.3987	1.5421	2220.624	810527.76	17.25%
900	9.466	7.8707	1.5953	2297.232	838489.68	16.85%
1000	9.896	8.392	1.504	2165.76	790502.4	15.20%

טבלה-2

server	N					
	cost k=0	cost k*	diff	diff for day	diff per year	diff ratio
100	11.23	4.2	7.03	10123.2	3694968	62.60%
200	22.46	5.51	16.95	24408	8908920	75.47%
300	33.508	6.567	26.941	38795.04	14160189.6	80.40%
400	44.892	7.488	37.404	53861.76	19659542.4	83.32%
500	56.155	8.013	48.142	69324.48	25303435.2	85.73%
600	67.665	8.652	59.013	84978.72	31017232.8	87.21%
700	78.326	9.199	69.127	99542.88	36333151.2	88.26%
800	90.125	9.729	80.396	115770.24	42256137.6	89.20%
900	100.909	10.128	90.781	130724.64	47714493.6	89.96%
1000	111.5985	10.648	100.9505	145368.72	53059582.8	90.46%

טבלה 3

server	N ²					
	cost k=0	cost k*	diff	diff for day	diff per year	diff ratio
100	937.889	12.734	925.155	799333.92	291756880.8	98.64%
200	3934.56	13.976	3920.584	3387384.58	1236395370	99.64%
300	9022.25	14.9	9007.35	7782350.4	2840557896	99.83%
400	16326.6	15.869	16310.731	14092471.6	5143752128	99.90%
500	25770	16.15	25753.85	22251326.4	8121734136	99.94%
600	37516	16.355	37499.645	32399693.3	11825888047	99.96%
700	50990.42	17.36	50973.06	44040723.8	16074864202	99.97%
800	67333.99	17.924	67316.066	58161081	21228794574	99.97%
900	85142.99	17.946	85125.044	73548038	26845033876	99.98%
1000	105029.988	18.246	105011.7416	90730144.7	33116502831	99.98%

טבלה 4

הנחות -

1. יחידת הזמן המדוברת היא דקה.
2. המוקד המדובר עובד 24 שעות ביממה, 365 ימים בשנה.

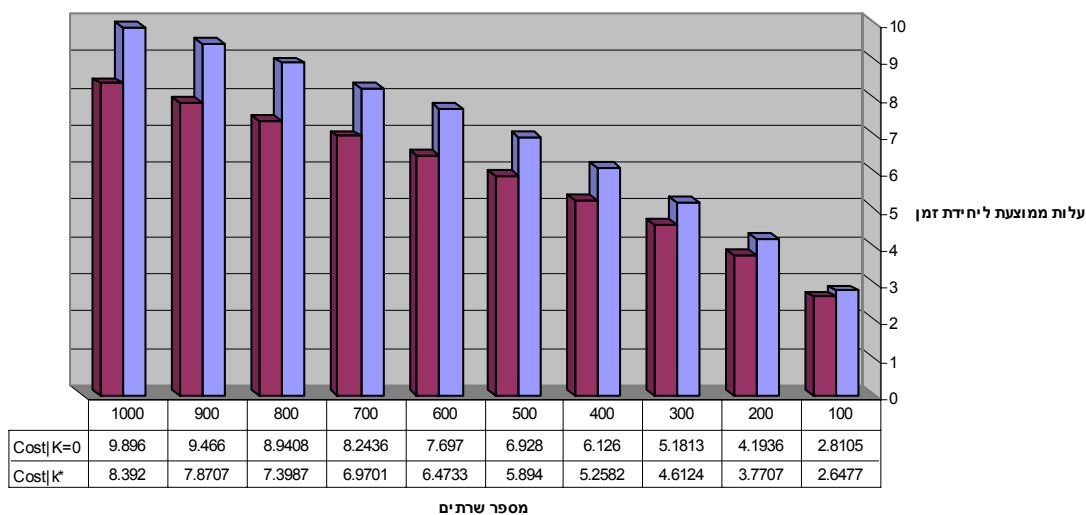
הסברים לעמודות המובאות בטבלאות -

- **cost|k=0** - העלות הכללית ביחידות כסף שנמצאה ללא השימוש בסף, לכל גודל מערכת שנבדק ליחידת זמן.
- **cost|k*** - העלות הכללית ביחידות כסף שנמצאה עבור הסף האופטימאלי, לכל גודל מערכת שנבדק, ליחידת זמן.
- **diff** - ההפרש בין שני האחרונים, שהוא בעצם החיסכון שמקנה הסף ליחידת זמן.
- **diff for day** - ההפרש מחושב עבור מחזור של יום שלם.
- **diff per year** - ההפרש המחושב עבור שנה.
- **diff ratio** - $\text{Diff}/(\text{cost}|k=0)$, כאמור אחוז החיסכון בין העלות כאשר משתמשים בסף לבין העלות כאשר לא הסף.

מטבלאות 2-4 ניתן לראות בקלות עד כמה השימוש בסף מוריד את עלויות המערכת הנובעות כתוצאה מזמני המתנה וכמה כסף נחסך במונחים של יחידת זמן, יום, ושנה.

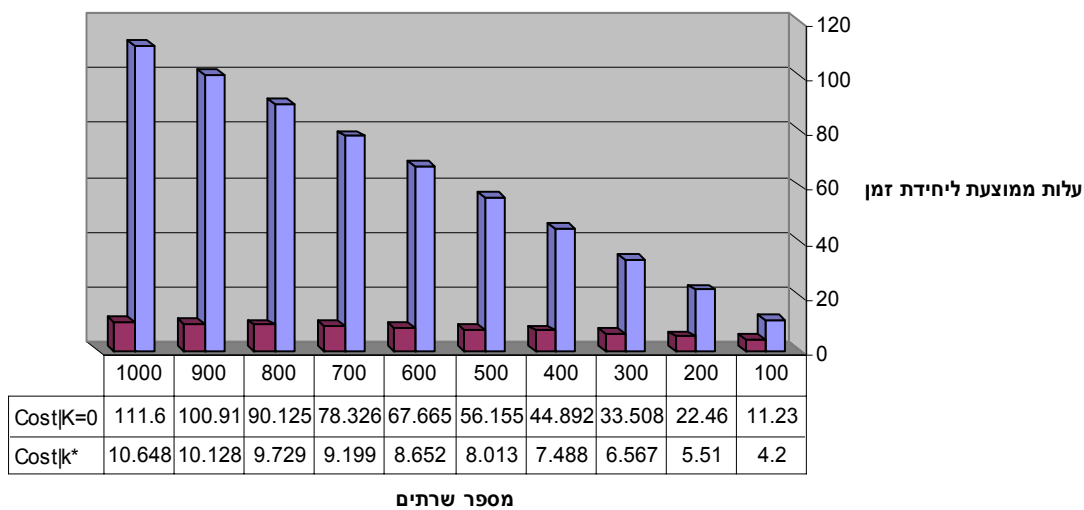
בתרשימים 16-18 מוצגים הנתונים הנ"ל עבור היחסים $\sqrt{N}$, N , N^2 , , במונחים של יחידות זמן.

עלויות עם ובלי סף עבור יחס שורש



Cost|K=0 Cost|k*

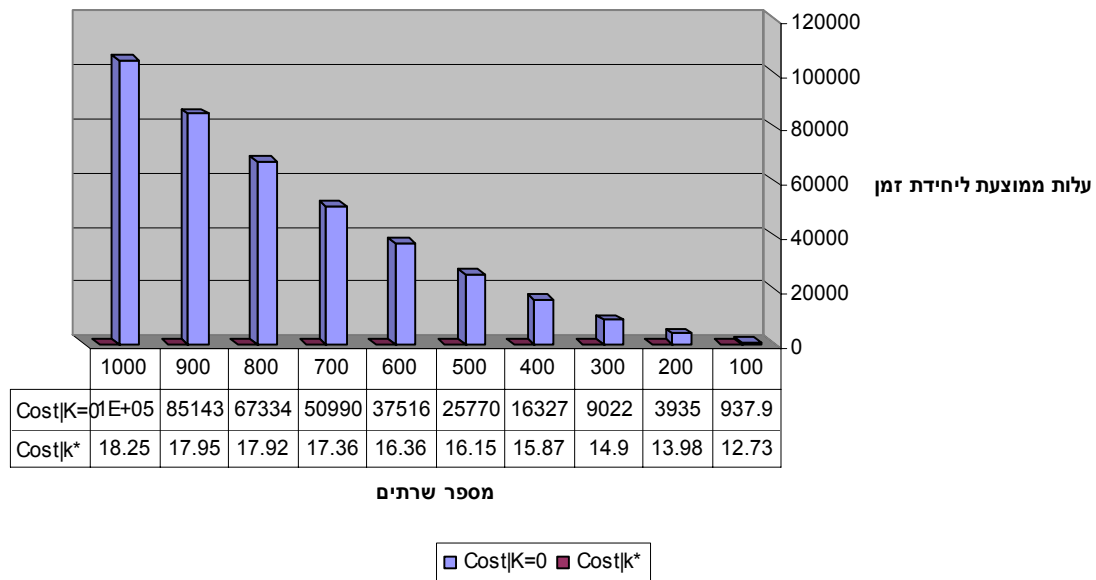
תרשים 16
עלויות עם ובלי סף עבור יחס N



Cost|K=0 Cost|k*

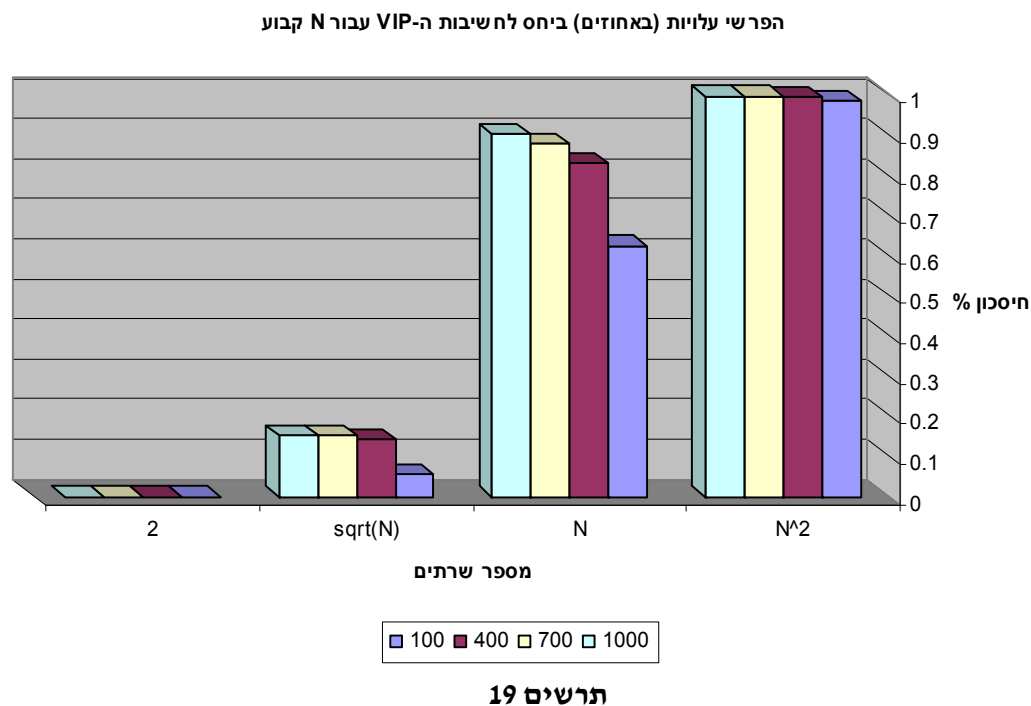
תרשים 17

עלויות עם ובלי סף עבור יחס N^2



תרשים 18

בתרשים 19 מודגם כיצד אחוז החיסכון משתנה כאשר יחס עלויות ההמתנה של הלקוחות השונים קבוע, ביחס לגודל המערכת, קרי – מספר השרתים. לשם הנוחות בחרנו להציג את התרשים הנייל עבור $N=100,400,700,1000$.



מהגרף המוצג בתרשים 19 ניתן הלבחין בשתי מגמות עיקריות:

1. ליחס עלות ההמתנה של הלקוחות ה-VIP ביחס ללקוחות הרגילים יש השפעה מכריעה על מידת היעילות שבשימוש באסטרטגיית הסף. שוב אנו מבחינים שעבור יחס עלויות של 2 אין כמעט חסכון בשימוש בסף וזאת מכיוון שכפי שראינו בעקומות הקודמות שהוצגו, הסף האופטימאלי המתקבל במרבית המקרים הוא אפס. אולם עבור יחסי עלויות של N ו- N^2 אנו מוצאים חיסכון של כמעט 100%.
2. קיימת כאן מגמה ברורה גם לגבי גודל המוקדים- עבור יחס קבוע של עלויות המתנה, ככל שהמוקד גדול יותר כך החיסכון גדל. למשל עבור יחס עלויות המתנה של N ניתן לראות כי עבור מוקד המעסיק 100 מוקדנים החיסכון הוא כ-60% ואילו עבור מוקד שמעסיק 1,000 מוקדנים החיסכון מתקרב ל-90%.

6.2 מודל N

6.2.1 הגדרת המודל והתסריט הנבדק -

הקדמה -

במוקדי שירות רבים ישנם מקרים בו לפלח לקוחות מסוים ישנה עדיפות על לקוחות אחרים. דוגמא לכך היא מוקד המסווג לקוחות ע"פ ערך חיי לקוח. בד"כ הסיווג המרכזי הוא לארבעה פלחי לקוחות – לקוחות פלטינה וזהב – שהם הלקוחות בעלי ערך הלקוח הגבוה ביותר, ולקוחות ברזל ועופרת – שהם הלקוחות בעלי הערך הנמוך ביותר. תרשים 1 מציג את פירמידת הלקוחות הנהוגה ב-CRM:

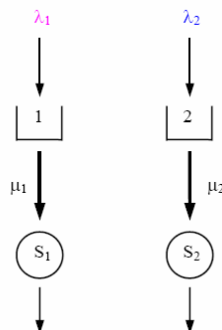


תרשים 1

מתוך גישה זו נוצר הצורך להטיב את השירות עם הלקוחות הערכיים יותר – לקוחות VIP לעומת הלקוחות הפחות ערכיים – הלקוחות הרגילים.

מקרה נוסף המתאים למודל N הוא מקרה בו ישנה דרישה לשני שירותים שונים, לדוגמא – במוקד בנקאי ישנם לקוחות המבקשים שירותים כלליים, ולקוחות המבקשים שירותים הנוגעים לניירות ערך.

מודל N – כפי שראינו בפרק 4.2, המאמר מציג את מודל I^2 שמתאר שתי מערכות מקבילות, אשר את כל אחת מהן ניתן לתאר מערכת $M/M/N$. מודל I^2 מתואר בתרשים 2.



תרשים 2

כל שרת מתמחה בשירות מסוים ויחיד ומשרת לקוחות המבקשים את שירות זה בלבד (ההבדלה של השירותים היא לא בהכרח בסוג השירות, אלא יכולה להיות ברמת השירות). נניח כי לקוחות מסוג אחד הם לקוחות הליבה של המערכת. הם נחשבים כלקוחות חשובים יותר מלקוחות אחרים ולכן ישורתו על ידי שרת מומחה – שרת 1.

ניתן לנתח את המערכת כשתי מערכות $M/M/N$, עבור הנתונים הבאים:

$$\lambda_1 = 1.3 \cdot \rho \cdot N$$

$$\lambda_2 = 0.4 \cdot \rho \cdot N$$

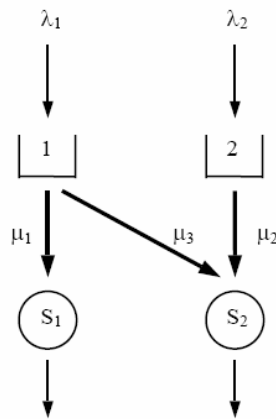
$$\mu_1 = \mu_2 = 1$$

$$S_1 = S_2 = N$$

כאשר ρ הוא עומס התנועה במערכת (מספר דקות השירות הנכנסות למערכת לדקה).

בטבלה 1 בסעיף 4.2 ניתן לראות כי בעומס גבוה השרת הראשון עמוס מאוד (נצילות של 97.5%), ולכן נוצר מצב כי הלקוחות מסוג 1 נאלצים להמתין זמנים ארוכים בתור, בעוד שהעומס על השרת השני אינו גדול כלל (30%).

הפתרון המוצג לזה הוא מודל N , המתואר בתרשים 3.



תרשים 3

מודל זה מנצל את התוצאה ששרת 2 אינו עסוק כמו שרת 1 ולכן הוא יוכל לעזור לשרת 1 ע"י מתן שירות ללקוחות מסוג 1. מודל זה מצריך הכשרה נוספת של שרת 2 על מנת שיוכל לשרת לקוחות מסוג 1. הנחה הגיונית במקרה זה היא שהשירות שניתן ללקוחות 1 ע"י שרת 2 יהיה איטי יותר.

הגדרת הניתוב –

תחת מודל זה קיימים תסריטים אפשריים רבים, כאשר עבודה זו תתמקד בתסריט אחד. התסריט המדובר עונה על שתי שאלות –

1. מה קורה כאשר מגיע לקוח?

- אם הלקוח הוא לקוח VIP אזי –

0 אם קיימים שרתים פנויים ב-S1 הלקוח מטופל ע"י שרת פנוי רנדומאלי.

0 אחרת, אם קיימים שרתים פנויים ב-S2, הלקוח מטופל ע"י שרת פנוי

רנדומאלי מ-S2.

0 אחרת, הלקוח יצטרף לתור אחד.

- אם הלקוח הוא לקוח רגיל אזי -

0 אם קיימים שרתים פנויים ב-S2, ותור אחד ריק, הלקוח יקבל שירות ע"י

שרת פנוי רנדומאלי מ-S2.

0 אחרת, הלקוח יצטרף לתור שתיים.

2. מה קורה כאשר שרת מתפנה?

- אם השרת המתפנה הוא מ-S1.

0 השרת יבדוק אם יש לקוחות בתור אחד, אם כן הוא ישרת את הלקוח

הבא בתור על פי מדיניות FCFS.

0 אחרת, השרת יעבור למצב Idle.

- אם השרת הפנוי הוא מ-S2.

0 השרת יבדוק אם יש לקוחות בתור אחד, אם כן הוא ישרת את הלקוח

הבא בתור על פי מדיניות FCFS.

0 אחרת, השרת יבדוק אם יש לקוחות בתור שתיים, אם כן הוא ישרת את

הלקוח הבא בתור על פי מדיניות FCFS.

0 אחרת, השרת יעבור למצב Idle.

תסריט זה מתאר מצב בו ללקוחות מסוג אחד ישנה עדיפות מוחלטת על לקוחות מסוג שתיים, שכן לקוחות אלו חשובים יותר לארגון ונרצה לתת להם את השירות הטוב ביותר.

בעייתיות התסריט –

כפי שניתן לראות בסעיף 4.2, עבור מקרה בו קיים שרת אחד מכל סוג, כבר בעומס של 0.87 המערכת מתפוצצת – כלומר, העומס על השרת בקבוצה S2, גבוה מהעומס בו הוא יכול לטפל. מצב זה נגרם מ"עזרה" יתרה שנותן השרת ב-S2 לשרת ב-S1. מטרת הסימולציה שנערכה בפרויקט זה (תתואר בהמשך), היא להראות את השפעת מספר השרתים על התנהגות המערכת. כלומר – עבור עומס קבוע, קיים N (מספר השרתים) עבורו המערכת תגיע במצב יציב. לדוגמא – עבור עומס של 0.9 קיים N גדול דיו כך שנקודת ההתפוצצות של המערכת תשאף לאחד.

6.2.2 הגדרת הסימולציה.

הנחות הסימולציה:

- המופע הוא תהליך פואסוני הומוגני בזמן עם הפרמטרים λ_1 ו- λ_2 (עבור לקוחות VIP ולקוחות רגילים בהתאמה).
- השירות הוא מ"מ אקספוננציאלי ב"ת בהגעות עם פרמטרים הבאים:
 $0 \leq \mu_1$ - קצב השירות של שרתים בקבוצה S1 ללקוחות VIP.
 $0 \leq \mu_2$ - קצב השירות של שרתים בקבוצה S2 ללקוחות רגילים.
 $0 \leq \mu_3$ - קצב השירות של שרתים בקבוצה S2 ללקוחות VIP.
- אין נטישות של הלקוחות או חסימה של שיחות במערכת. במקרה בו קיימות נטישות או חסימות המערכת תמיד יציבה וכן המקרה לא רלוונטי.
- הנתונים לסימולציה נלקחו מתוך Introduction to SBR and it's operational Complexities (גרנט ומנדלבאום, 2003). והם:

$$\lambda_1 = 1.3 \cdot \rho \cdot N$$

$$\lambda_2 = 0.4 \cdot \rho \cdot N$$

$$\mu_1 = \mu_2 = 1$$

$$\mu_3 = 0.5$$

$$S_1 = S_2 = N$$

(הנתונים מתייחסים למודל N הבסיסי המוצג בתרשים 3.)

הנחות אלו הוכחו כסבירות במקרה של מודל שירות טלפוני.

מאחר וכל הזמנים שהוגרלו נלקחו מההתפלגות האקספוננציאלית, השתמשנו בתכונת חוסר הזיכרון של ההתפלגות, ובכל איטרציה הוגרלו ערכי הזמן של שלושה אירועים :

1. הזמן עד הגעת הלקוח VIP הבא.
 2. הזמן עד הגעת הלקוח הרגיל הבא
 3. הזמן עד ששרת מקבוצת שרתים S2 מתפנה משירות לקוח VIP.
 4. הזמן עד ששרת מקבוצת שרתים S2 מתפנה משירות לקוח רגיל.
 5. הזמן עד ששרת מקבוצת שרתים S1 מתפנה משירות לקוח VIP.
- "האירוע הבא" מוגדר להיות המינימום מבין שלושת האירועים הללו, והזמן המתאים מתווסף לזמן הנוכחי המוגדר בסימולציה ומתקדם לאורך הריצה.

מטרת הסימולציה:

הסימולציה נועדה לאשש את ההנחה כי עבור מספר שרתים גדול דיו המערכת תגיע למצב יציב עבור כל עומס שהוא. מהמאמר המוצג ב-4.2 ידוע כי הבעיה נובעת מהעומס על השרתים בקבוצה S2, שנוצר בעקבות עודף העזרה שהם נותנים לשרתים בקבוצה S1. מכאן מטרת הסימולציה היא למצוא את העומס על השרתים בקבוצה S2 כפונקציה של מספר השרתים.

מבנה הסימולציה:

- הסימולציה נכתבה בשפת C++ . הקוד מובא בנספח 4.
- הסימולציה עובדת בהתאם לכללי הניתוב בשוהגדו בסעיף 6.1.1
- הסימולציה מדמה את המודל המדובר, תחת ההנחות שהוזכרו, עבור כל עומס בין 0.02 ל-1 (בקפיצות של 0.01).
- הסימולציה אוגרת נתונים שונים, על מנת לחשב את הפלטים שיתוארו בהמשך.
- כל סימולציה הורצה עבור 20,000 יחידות זמן.
- הסימולציה הורצה עבור מספר מערכים של שרתים לפי הפירוט הבא :
 - 0 1-10 שרתים מכל סוג, בקפיצות של אחד.
 - 0 10-100 שרתים מכל סוג, בקפיצות של עשרה.
 - 0 200, 400, 600 שרתים מכל סוג.

נכונות האלגוריתם :

- על מנת לבדוק את נכונות הסימולציה הרצויה אותה עבור $S1=1$ ו- $S2=1$. מקרה זה זהה למקרה המדובר בסעיף 4.2 והשוואה של התוצאות שקיבלנו לתוצאות שהתקבלו בסימולציה המתוארת ב-4.2 יצאו זהות.
- בדיקה נוספת נעשתה ע"י איפוס של מספר השרתים ב- $S1$ ומכאן המודל זהה למודל V עם סף אפס. השונו בין תוצאות הסימולציה בשני המקרים ומצאנו כי הן אכן זהות.

דיוק:

על מנת להגיע לרמת דיוק גבוהה הרצנו 50 רפליקציות של הסימולציה ואמדנו את התוצאות המבוקשות בעזרת ממוצע.

הפלט לסימולציה:

כל הפלטים הועברו לגיליון excel בפורמט הבא :

total_reg	Total VIP	RegS2	VipS2	VipS1	row
-----------	-----------	-------	-------	-------	-----

הסבר לגבי העמודות :

- row – העומס על המערכת, כאמור נע בין 0.02 ל-1 בקפיצות של 0.01.
- VipS1 – מספר לקוחות ה- VIP אשר שורתו ע"י שרתים מקבוצה $S1$.
- VipS2 – מספר לקוחות ה- VIP אשר שורתו ע"י שרתים מקבוצה $S2$.
- RegS2 – מספר הלקוחות הרגילים אשר שורתו ע"י שרתים מקבוצה $S2$.
- Total VIP – סה"כ לקוחות ה- VIP שהגיעו למערכת.
- Total Reg – סה"כ הלקוחות הרגילים שהגיעו למערכת.

מתוך נתונים אלו חושב העומס הממוצע על שרת מקבוצה $S2$ באופן הבא :

$$\frac{Vip_S2}{run_time} + \frac{total_reg}{run_time} = \frac{Vip_S2}{0.5 \cdot N_2} + \frac{total_reg}{N_2}$$

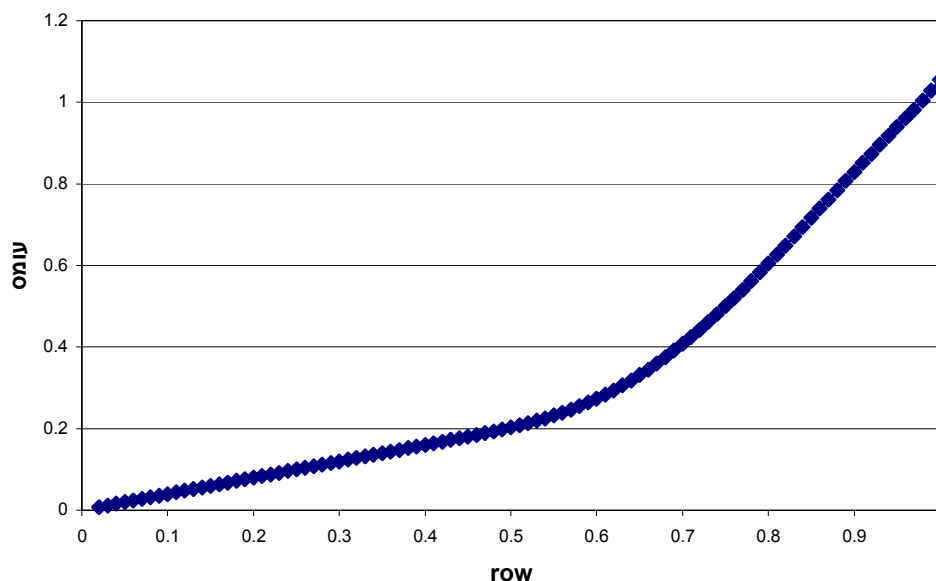
הסבר לחישוב:

הנוסחא מתבססת על חישוב העומס כמכפלת קצב ההגעה בזמן השירות הממוצע, כל זאת לחלק למספר השרתים. קצב ההגעה חושב ע"י חלוקת מספר הלקוחות המתאים שהגיע בכל הסימולציה לחלק בזמן הסימולציה.

העומס על השרתים מקבוצה S2 מורכב מהעומס שנוצר כתוצאה משירות לקוחות ה-VIP וכן מהעומס הנוצר כתוצאה משירות הלקוחות הרגילים, מכאן:
העומס על שרתים מקבוצה S2 = העומס שנוצר כתוצאה משירות לקוחות ה-VIP + העומס הנוצר כתוצאה משירות הלקוחות הרגילים.

6.2.3. תוצאות הסימולציה:

עבור כל סט של שרתים הופק הגרף המתאר את העומס הממוצע על שרת מקבוצת השרתים S2 כפונקציה של העומס הכללי על המערכת, על פי החישוב שהוצג בסעיף הקודם. דוגמא לגרף זה עבור 40 שרתים מכל סוג ניתן לראות בתרשים 5:



תרשים 5

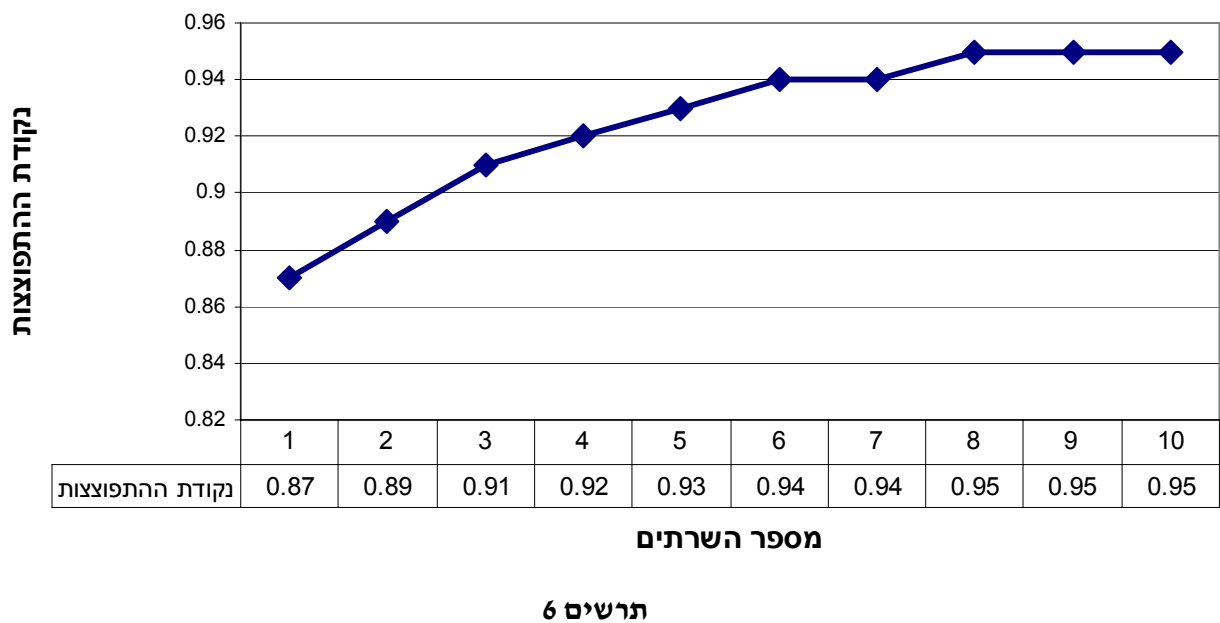
ציר X – row - עומס התנועה במערכת כולה (מספר דקות השירות הנכנסות למערכת לדקה).

ציר Y – העומס הממוצע על שרת מקבוצת שרתים S2.

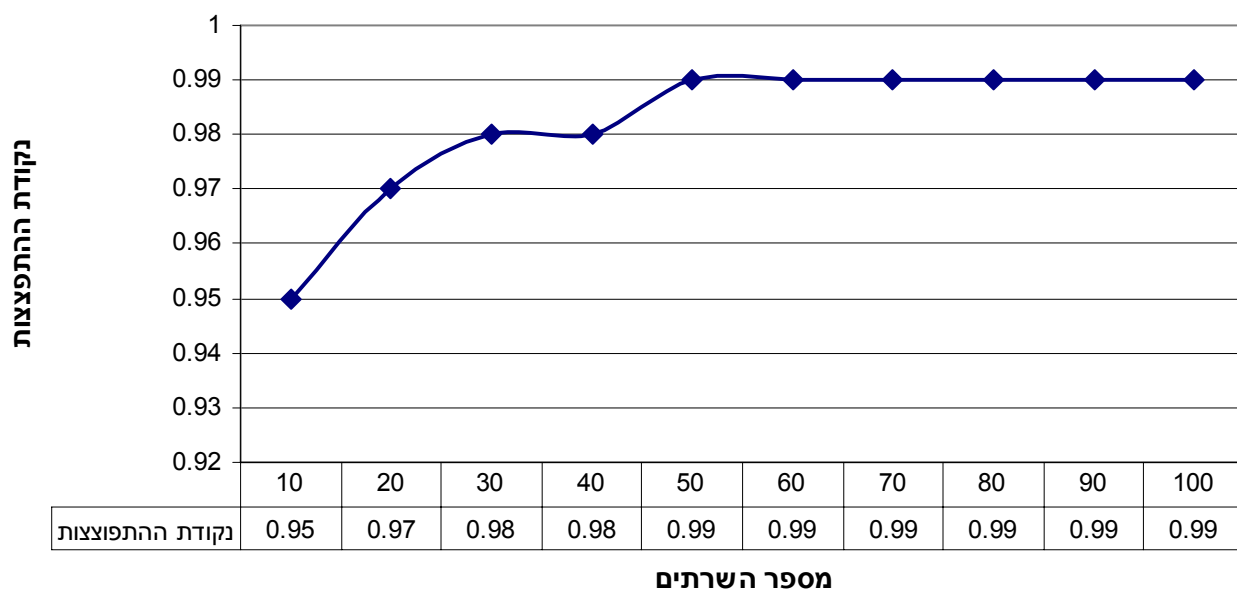
הגרפים המלאים מצורפים ב-CD ROM

כתוצאה מהגרף ומגיליונות העבודה הבונים אותו, ניתן למצוא את נקודת ההתפוצצות – העומס על המערכת עבורו העומס הממוצע על שרת מקבוצה S2 גדול מ-1. לדוגמא בתרשים זה ניתן לראות כי כאשר העומס על המערכת הוא 0.98 העומס על השרת השני הוא 1.003. כלומר למרות שהעומס על המערכת אינו מקסימאלי, העומס הממוצע על שרת מקבוצה S2 גבוה מ-1, ולכן השרתים לא מסוגלים לטפל בכל הלקוחות.

הגרף בתרשים 6 מסכם את נקודות ההתפוצצות עבור N שרתים, $N=1, \dots, 10$:

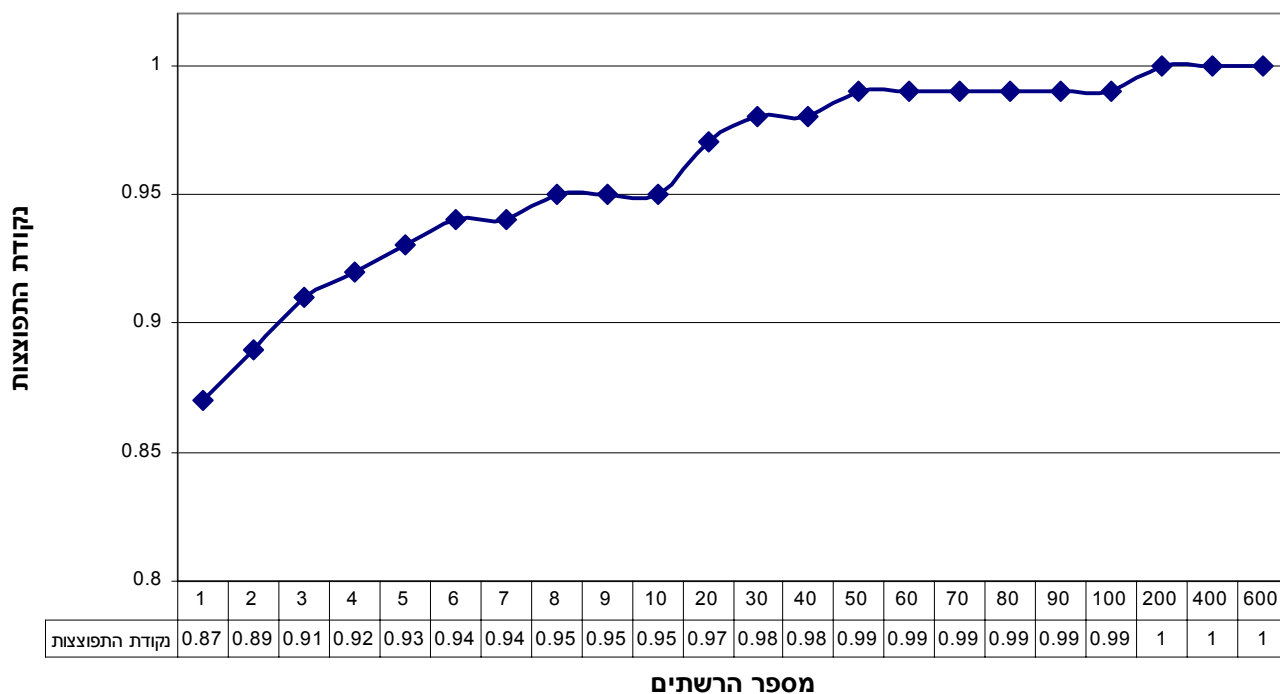


הגרף בתרשים 7 מסכם את נקודות ההתפוצצות עבור $N=10,20,\dots,100$ שרתים:



תרשים 7

הגרף בתרשים 8 מסכם את נקודות ההתפוצצות עבור כל השרתים שהוצגו במבנה הסימולציה:



תרשים 8

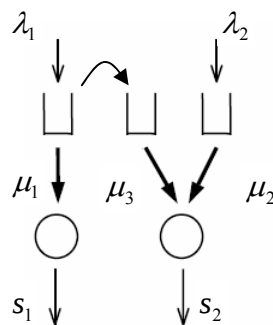
6.2.4. ניתוח התוצאות:

ע"פ ההסבר התיאורטי שהבאנו בסעיף 1.5 הוכח כי עבור כל עומס מסוים על המערכת (row), ניתן למצוא N (מספר שרתים) גדול דיו על מנת שהמערכת תגיע למצב יציב. מכאן נובע כי במערכות גדולות, קרי – מערכות שירות עם מספר גדול של שרתים, המערכת תמיד תהייה יציבה. תוצאות הסימולציה שהוצגו בסעיף 3 מאששות גם הן את טענה זו. מהתוצאות ניתן לראות כי ככל שמספר השרתים גדל, נקודות ההתפוצצות שואפת לאחד. לדוגמא, עבור 200 שרתים מכל סוג, המערכת מתפוצצת באחד – כמו במצב הטבעי של מערכות סטוכסטיות. תופעה זו מתארת למעשה את היתרון לגודל – ככל שמספר השרתים גדול יותר ביצועי המערכת יהיו טובים יותר. לדוגמא – זמן ההמתנה תחת אותו העומס על המערכת, כאשר המערכת גדולה יותר, יהיה קצר יותר ממערכת בה יש מספר קטן של שרתים. חשוב להדגיש כי הגדלת מספר השרתים אינה פותרת את בעיית חוסר האיזון בין נצילות השרתים כפי שהוסבר בפרק 4.2.

ריבוי השרתים אכן מביא לידי ביטוי את היתרון לגודל בכך שנקודת ההתפוצצות מתרחשת באופן טבעי יותר משום שכפי שראינו בפרק 4.2, השרתים בקבוצה S_2 עובדים בנצילות גבוהה יותר ולכן הטיפול בלקוחות הרגילים טוב יותר. בעצם נוצר מצב בו אנו מטשטשים את בעיית האיזון ולא פותרים אותה. כדי לפתור את בעיה זו ניתן לבדוק מקרים נוספים כמו הוספת מדיניות סף כלשהי.

הסבר תיאורטי לתוצאות שהתקבלו –

נסתכל על שתי מערכות – המערכת הראשונה היא מודל N שהוצג בתרשים 3, והמערכת השנייה המוצגת בתרשים 4:



תרשים 4

המערכת המתוארת בתרשים 4 מורכבת למעשה ממערכת $M/M/N/N$ – כלומר מערכת ללא תור, ומערכת הניתן לתאר כמודל V . לקוחות מסוג אחד אשר מגיעים למערכת ומוצאים אותה מלאה (כלומר, כל N השרתים עסוקים) עובר למערכת השנייה (מערכת ה- V).

ניתן להגיד כי התור של הלקוחות מסוג 2 במערכת ה-N המקורית לא יהיה גדול מהתור של הלקוחות מסוג 2 במערכת השנייה. הסיבה לכך היא שקצב הגעת הלקוחות מסוג אחד לקבוצת שרתים S2 במערכת הראשונה יהיה נמוך בהרבה מקצב הגעת הלקוחות מסוג אחד לקבוצת שרתים S2 במערכת השנייה שכן במערכת הראשונה קיים תור במאכלס לקוחות מסוג אחד ובמערכת השנייה לא.

מכאן, אם נוכיח כי המערכת השנייה מגיעה למצב יציב, נוכל להגיד כי גם המערכת הראשונה מגיעה למצב יציב (שכן הראנו כי הבעייתיות נוצרת מהעומס על השרתים ב-S2).

נגדיר את קצב הגעת הלקוחות מסוג אחד לקבוצת שרתים S2 במערכת השנייה כ- $\pi_1 = P\{block\} \cdot \lambda_1$. כאשר $P\{block\}$ הוא אחוז הלקוחות המגיעים למערכת ה-M/M/N/N ומוצאים את המערכת מלאה. על פי PASTA פרופורציה זו שקולה לפרופורציה הזמן שכל השרתים במערכת עסוקים.

מכאן העומס על קבוצת השרתים S2 במערכת השנייה הוא :

$$\frac{\pi_1}{\mu_3} + \frac{\lambda_2}{\mu_2} = \frac{\lambda_1 \cdot P\{block\}}{\mu_3} + \frac{\lambda_2}{\mu_2}$$

ולכן על מנת לדרוש כי המערכת במצב יציב נדרוש כי :

$$\frac{\lambda_1 \cdot P\{block\}}{\mu_3} + \frac{\lambda_2}{\mu_2} < N_2$$

ואת זה נרצה להוכיח.

ע"פ המאמר המוצג בסעיף 4.2, המערכת המקורית נבנתה כך שהעומס על השרתים ב-S2 יביא את המערכת למצב יציב ולכן ידוע כי :

$$\frac{\lambda_1 - N_1 \mu_1}{\mu_3} + \frac{\lambda_2}{\mu_2} \leq N_2$$

כיוון שהגדרנו כי מטרת הסימולציה היא למצוא את מספר השרתים אשר יביא את המערכת למצב יציב עבוד עומס קבוע נניח כי העומס קטן בהחלט מאחד ולכן :

$$\frac{\lambda_1 - N_1 \mu_1}{\mu_3} + \frac{\lambda_2}{\mu_2} < N_2$$

לכן כל שנותר לנו להוכיח הוא שעבור N גדול דיו מתקיים כי :

$$\frac{\lambda_1 \cdot P\{block\}}{\lambda_1 - N_1 \mu_1} \xrightarrow{N \uparrow \infty} 1$$

הסבר אינטואיטיבי לגבול זה הוא שבמערכת עם עומס גבוה המערכת תעבוד בצורה כמעט דטרמיניסטית ולכן תשחרר $N_1 \mu$ לקוחות ליחידת זמן. מכיוון שכמות הלקוחות שהגיעו למערכת הייתה λ_1 , כמות הלקוחות שלא יכנסו למערכת היא בעצם $\lambda_1 - N_1 \mu_1$, כאשר זהו בדיוק מספר הלקוחות שייחסמו במערכת ($\lambda_1 \cdot P\{block\}$). טענה זו הוכחה על ידי Bell and Williams במאמר "Dynamic Scheduling of a System with two Parallel Servers in Heavy Traffic With Complete Resource Pooling: Asymptotic Optimality of a Continuous Review Threshold Policy"

6.2.5 הצעות למחקר המשך :

מודל N הוא מודל שנחקר במאמרים אקדמיים רבים. בעבודת התזה של מר איתי גורביץ' הוכח כי קיים פיתרון אופטימאלי עבור התסריט הבא :

נסתכל על המודל המוצג בתרשים 3. המקרה המדובר הוא מקרה בו $\mu_1 \geq \mu_2 = \mu_3$. ידוע כי הלקוחות המגיעים בקצב λ_2 הם לקוחות הליבה של המערכת (לקוחות ה-VIP).

נבחן את המוטיבציה לתיכון זה – לקוחות הליבה דורשים סוג מסוים של שירות אשר יכול להינתן אך ורק על ידי השרתים בקבוצה S2. הלקוחות המגיעים בקצב λ_1 הם לקוחות רגילים, אשר דורשים שירותים רגילים האורכים זמן קצר יותר. השרתים המנוסים בקבוצה S2 מסוגלים לטפל הן בשירותים הרגילים והן בשירות המיוחד. לעומת זאת, השרתים בקבוצה S1 מסוגלים לטפל אך ורק בשירותים הרגילים ואינם מוכשרים לטפל בלקוחות הליבה.

יהלום ומנדלבאום הוכיחו במאמר "Optimal Scheduling Of a Multi-Server Multi-Class Non Preemptive Queueing System" כי המדיניות הבאה היא מדיניות אופטימלית למקרה זה :

- כאשר שרת מקבוצת שרתים $S1$, מסיים לתת שירות הוא ישרת לקוח מלקוחות הליבה, אם כאלו ממתינים בתור.
- כאשר שרת מקבוצת שרתים $S2$, מסיים לתת שירות הוא ישרת לקוח מהלקוחות הרגילים, אם כאלו ממתינים בתור. אחרת, הוא ישרת לקוח מלקוחות הליבה, אם יש לפחות $q_1^*(I2)$ לקוחות ליבה הממתינים בתור. $q_1^*(I2)$ היא פונקציה לא עולה של $I2$ - מספר השרתים הפנויים בקבוצת שרתים $S2$.
- כאשר מגיע לקוח הלקוחות הליבה, והתור המתאים ריק, הוא יקבל שירות משרת פנוי מסוג $S2$, אם קיים כזה. אחרת הוא יצטרף לתור.
- כאשר מגיע לקוח מקבוצת הלקוחות הרגילים, ואין לקוחות בתור המתאים הוא יקבל שירות על ידי שרת פנוי מקבוצת השרתים $S1$, אם יש כזה. אחרת, אם ישנם לפחות $I2$ שרתים פנויים בקבוצת שרתים $S2$ (או אם $q_1^*(I2) \leq 1$) הוא יקבל שירות על ידי שרת מקבוצת השרתים $S2$.

קיימים תסריטים שונים אשר מבוססים על מודל N המוצג בתרשים 3. עד לסיום פרויקט זה לא נמצא פתרון אופטימאלי עבור כל התסריטים, וכן הוכח כי עבור תסריטים מסוימים לא קיים פתרון אופטימאלי. בפרויקטים עתידיים ניתן למצוא תסריטים נוספים להם קיים פתרון אופטימאלי, וכן את הפתרון המתאים.

פרק 7

סיכום

השימוש בניתוב מבוסס מיומנות קיים כיום במוקדים רבים בארץ. השימוש נהפך להכרחי עקב ההבדלים הרבים בין הלקוחות השונים ובין השירותים שהם דורשים. מוקדי שירות טלפוניים הם תוצאה ישירה של דרך החיים המתפתחת במאה ה-20. הצרכן הממוצע עובד יותר שעות, וכתוצאה מכך קטן הזמן הפנוי שברשותו. ככל שיורד שיעור הזמן הפנוי של הצרכן, כל המתנה או טרחה נתפסת כארוכה ומטרידה יותר.

במוקדי שירות טלפוני ניתן להגדיר את מדיניות ה- SBR (Skill-Based Routing) כיכולת לנתב שיחה אל הנציג בעל היכולת הגבוהה ביותר לענות לשיחה זו. SBR לוקח בחשבון את דרישות הלקוח הנובעות מסיבות שונות. ה- SBR מאפשרת לארגון לדבר עם הלקוחות בקול אחד, מאפשרת לשרתים להתפנות ולטפל באופן יותר פרודוקטיבי בלקוחות המועדפים ומעניקה כוח למנהלים להתמקד במשימות הליבה של ניהול משאבים, תהליכי למידה ותהליך איסוף המידע.

במוקד שירות טלפוני העובד ע"פ מודל V, כאשר קיימים שני סוגי לקוחות אשר לאחד מהם חשיבות רבה יותר מלשני יש טעם בשימוש במדיניות הסף שהוצגה בפרק 6.1. השימוש יביא לחיסכון בעלויות במקרה בו זמן ההמתנה של לקוחות ה-VIP (הלקוחות החשובים) יקר הרבה יותר מזמן ההמתנה של הלקוחות הרגילים. מבחינה פרקטית, מנהל האחראי על מוקד שירות טלפוני הדומה לזה המתואר כאן, יכול בקלות להגדיר לעצמו, על ידי קביעת הסף המתאים, את מידת העדיפות שיש לתת ללקוחות החשובים יותר כפונקציה של גודל המערכת עליה הוא אחראי- מספר השרתים העומדים לרשותו.

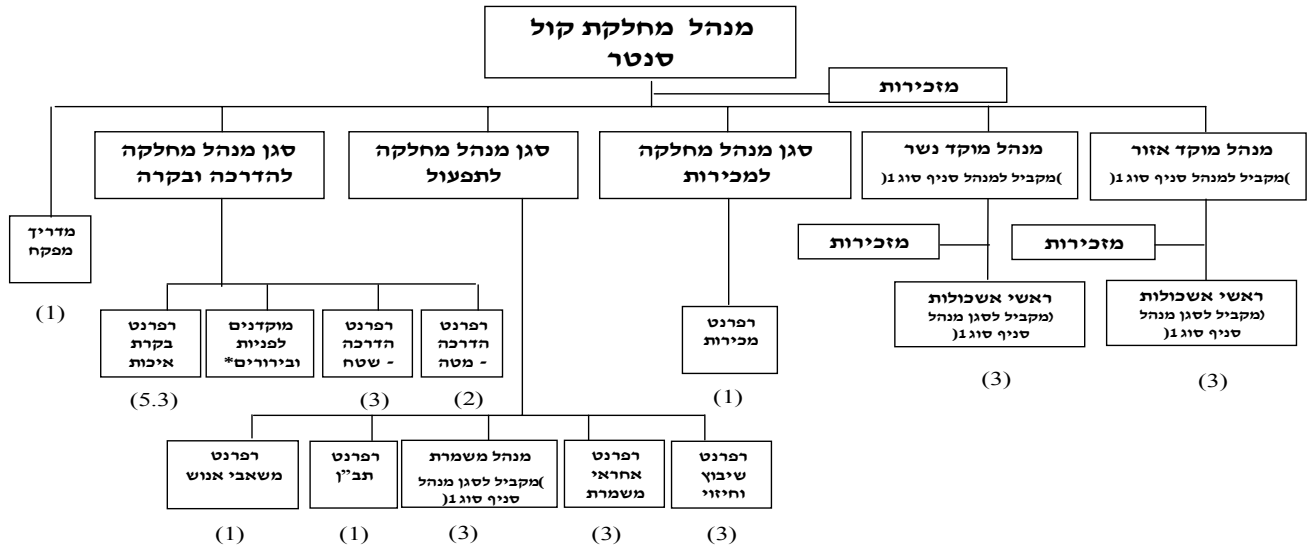
השימוש במודל N תחת התסריט שהוצג, במוקדי שירות טלפוני מותנה בגודל המוקד, כלומר – מספר השרתים במוקד. במקרה של מוקד קטן שימוש במודל N יביא להתפוצצות מהירה של המערכת. כלומר – כבר עבור עומס קטן על המערכת השרתים "העוזרים" לא יעמדו בעומס המוטל עליהם, הגדלת סדרי הגודל של המערכת מאפשרת לראות באופן ברור את היתרון לגודל.

במחקרים עתידיים ניתן לבדוק תרחישים שונים של מודל N אשר להם ניתן למצוא פתרון אופטימאלי. אפשרי לבדוק את יישום המודלים שהוצגו בפרויקט זה בפועל, תוך כדי התייחסות לשילובים שונים של המודלים ולהיעזר לצורך מחקר זה בנתונים ממוקדי שירות שונים.

1. Bell S.L., Williams R.J., "Dynamic Scheduling of a System with Two Parallel Servers in Heavy Traffic with Complete Resource Pooling: Asymptotic Optimality of a Continuous Review Threshold Policy", *Annals of Applied Probability*, Vol. 11, 2001.
2. Gans, N., Koole, G., Mandelbaum, A. "Telephone Call Center: Tutorial, Review and Research Prospects" (March, 2003) In <http://iew3.technion.ac.il/serveng/References>
3. Granett, O. and Mandelbaum, A. "An Introduction to Skills-Based Routing and its Operational Complexities" (May, 2003) In <http://iew3.technion.ac.il/serveng/References>
4. Halfin, S. and Whitt W. "QED Theorem " from QED Q's at http://iew3.technion.ac.il/~serveng/course2004B/Lectures/QED_Qs.pdf.
5. Maiane L'cuyer' Elix "Using Skill Based Routing To Enhance Contact Center Revenue And Performance" (May, 2003) In www.elixononline.com/reprts.asp?id=972.
6. Rodney, B. and Whitt, W. "Resouce Pooling And Staffing Call Centers With Skill Based Routing" (March, 2004) In <http://www.columbia.edu/~ww2040/recent.html>
7. "The Cultural Impact Skills Based Routing" In www.logicacmq.com

נספחים

נספח 1 – מודל ארגוני של פועלים ישיר.



נספח 2 – תיאורית ה-QED:

כאשר מתבוננים במערכת תורים מסוג $M/M/N$, ניתן לבחון שלוש נקודות מבט שונות:

- נקודת מבט של הלקוח – $0 < \alpha < 1$ $\lim_{n \rightarrow \infty} P_N \{Wait > 0\} = \alpha$
- נקודת המבט של המוקדן – $0 < \beta < \infty$ $\lim_{n \rightarrow \infty} \sqrt{N}(1 - \rho N) = \beta$
- נקודת המבט של מנהל המוקד – $N \approx R + \beta\sqrt{R}$
 $R = \lambda \cdot E(S) \quad large$

מכיוון ששלושת נקודות מבט אלו שקולות מתקיים:

$$\alpha = \left[1 + \frac{\beta\phi(\beta)}{\varphi(\beta)} \right]^{-1}$$

כאשר $\phi(\cdot), \varphi(\cdot)$ הן פונקציית הצפיפות / התפלגות של התפלגות נורמלית בטנדרטית בהתאמה.

במקרים הקיצוניים:

- אם $\alpha = 1 \Leftrightarrow \beta = 0$ אזי כל הלקוחות ימתינו ומכאן המוקד הוא מוקד מכוון יעילות. הדגש הוא על נצילות מקסימלית של העובדים ולא על השירות הניתן ללקוחות.
- אם $\alpha = 0 \Leftrightarrow \beta = \infty$ אזי אף לא אחד מהלקוחות ימתין, אך נצילות המערכת תהייה נמוכה. מכאן המוקד הוא מוקד מכוון איכות.

מדיניות QED היא מדיניות המתקבלת משילוב של מדיניות קיצונית אלו. (Quality and Efficiency Driven).

נספח 3 – קוד סימולציה עבור מודל V, מקרה בו קצבי ההגעה של לקוחות רגילים ו-VIP שווה:

```
#include<iostream.h>
#include<stdio.h>
#include<math.h>
#include<stdlib.h>
#include<fstream.h>

//function Min finds the minimum between three events
//arrivals of VIP customer, arrivals of Regular customer and
//the time the next customer leaves the service (which is Service time)
double Min(double,double,double);
int main()
{
    //miy is the service rate
    double miy=1;
    //betta is parmeter used in lambda calculation
    double betta=1;
    //total number of servers
    int N;
    //total simulation time. we took simulation time as 50,000 time units so the
    //system will arrive to her study
    double RunTime;

    cout << "Enter Total Number of servers" <<endl;
    cin >> N;
    cout << "Enter Total Time of Simulation" <<endl;
    cin >> RunTime;
```

```
//calculating lambda according to QED safty stuffing rule
```

```
double lambdaV=(N-beta*sqrt(N))/2;
```

```
double lambdaR=(N-beta*sqrt(N))/2;
```

```
ofstream Out1("data300.xls");
```

```
    if(!Out1)
```

```
    {
```

```
        cerr << "cannot open file" << '\t' << "datav300.xls" << '\t' << endl;
```

```
        exit(1);
```

```
    }
```

```
Out1<<"k"<<'\t'<<"AvgWaitTimeV"<<'\t'<<"AvgWaitTimeR"<<endl;
```

```
//initializing times
```

```
double Vip=0;
```

```
double Reg=0;
```

```
double Serv=0;
```

```
//initializing waitinig time - sum and avarage for each type of customers
```

```
double SumWaitTimeV;
```

```
double SumWaitTimeR;
```

```
double AvgWaitTimeV;
```

```
double AvgWaitTimeR;
```

```
//counter of number in queue for each customer type
```

```
int queueV;
```

```
int queueR;
```

```

//threshold for serving regular customer. serving rule -serving only VIP customers.
//if free servers is bigger than k the servers will serve regular customers

int k=0;
//number of free servers

int fs;
//counter of total customers

int totVcust;
int totRcust;
//current time

double now;

while(k<=(N/5))
{
    AvgWaitTimeV=0;
    AvgWaitTimeR=0;

    for (int rep=1;rep<6; rep++)
    {
        now=0;
        SumWaitTimeV=0;
        SumWaitTimeR=0;
        totVcust=0;
        totRcust=0;
        queueV=0;
        queueR=0 ;
        fs=N;
        cout<<"k="<<'\'t'<<k<<endl;
    }
}

```

```

while(now<RunTime)
{ double u=0,y=0,w=0,early_event=0;
  u=(rand()+1.)/(RAND_MAX+1.);
  Vip=(-log(u)/lambdaV);
  w=(rand()+1.)/(RAND_MAX+1.);
  Reg=(-log(w)/lambdaR);

  if(fs<N)
  {      y=(rand()+1.)/(RAND_MAX+1.);
        Serv=(-log(y)/(miy*(N-fs))); }
  else Serv=RunTime;
  early_event=Min(Vip,Reg,Serv);
  now+=early_event;

  if(now>100)
    { SumWaitTimeV+=early_event*queueV;
      SumWaitTimeR+=early_event*queueR;    }

  if(early_event==Serv)
    {      fs++;
      if(queueV>0)
        {queueV--;
          fs--; }
      else
        if(queueR>0 && fs>k)
          { queueR--;
            fs--; }
    }
}

```

```

if(early_event==Vip)
{
    if(fs>0) fs--;
    else queueV++;
    if(now>100) totVcust++;
}

if(early_event==Reg)
{
    if(fs>k()) fs--;
    else queueR++;
    if(now>100) totRcust++;
}
}

AvgWaitTimeV=AvgWaitTimeV+SumWaitTimeV/totVcust;
AvgWaitTimeR=AvgWaitTimeR+SumWaitTimeR/totRcust;

}

AvgWaitTimeV=AvgWaitTimeV/5;
AvgWaitTimeR=AvgWaitTimeR/5;

Out1<<k<<'\\t'<<AvgWaitTimeV<<'\\t'<<AvgWaitTimeR<<endl;
K++;
}

return 0;
}

```

```
double Min(double Vip,double Reg,double Serv)
{
    if(Vip>Reg)
        if(Reg>Serv) return Serv;
        else return Reg;
    else
        if(Vip>Serv) return Serv;
        else return Vip;
}
```

```
#include<iostream.h>
#include<stdio.h>
#include<math.h>
#include<stdlib.h>
#include<fstream.h>

double Min(double,double,double,double,double);

int main()
{
    // data
    // service times
    double miy1=1;
    double miy2=1;
    double miy3=0.5;
    //arrivals rate
    double lambda1;
    double lambda2;
    //number of servers
    int S1;
    int S2;
    //events times
    double VipArv;
    double RegArv;
    double VipServ;
    double RegServ;
    double VipHelp;
```

```

//queue
int VipQ=0;
int RegQ=0;
//fraction of VIP served by S2
int total_VIP=0;
int total_reg=0;
int VipS2=0;
int VipS1=0;
int RegS2=0;
//number of s2 serving VIP
int serv2vip=0;
int serv2reg=0;
// general variables
int FreeS1;
int FreeS2;
double row;
double now;
double RunTime=5000;
int replication;
int N_of_replication=10;
cout << "Enter number of VIP servers:" <<endl;
cin >> S1;
cout << "Enter number of REG servers:" <<endl;
cin>>S2;
cout << "Enter your RunTime:" <<endl;
cin >> RunTime;

```

```

ofstream Out1("data1.xls");

    if(!Out1)
    {
        cerr << "cannot open file" << '\t' << "data1.xls" << '\t' << endl;
        exit(1);    }

Out1<<endl;

Out1<<S1<<'\t'<<"number of S1="<<'\t'<<S2<<'\t'<<"number of S2="<<'\t'<<'\t'<<RunTime<<'\t'<<"Run Time="<<endl;
Out1<<"row"<<'\t'<<"VipS1"<<'\t'<<"VipS2"<<'\t'<<"RegS2"<<'\t'<<"Total VIP"<<'\t'<<"total_reg"<<endl;

for (row=0.02; row<1.01; row=row+0.01)
{
    int VipS1_replication=0;
    int VipS2_replication=0;
    int RegS2_replication=0;
    int total_VIP_replication=0;
    int total_REG_replication=0;

    for (replication=1 ; replication<=N_of_replication; replication++)
    {
        cout<<"row"<<'\t'<<row<<'\t'<<"replication"<<'\t'<<replication<<'\t'<<"S1 ="<<'\t'<<S1<<'\t'<<"S2 ="<<'\t'<<S2<<endl;

        VipQ=0;           //number VIP in queue
        RegQ=0;           //number VIP in queue
        total_VIP=0; //total number of VIP
        total_reg=0; //total number of reg
        serv2vip=0;       //number of server 2 serving VIP costomer
        serv2reg=0;       //number of server 2 serving VIP costomer
        VipS2=0;          //number of VIP costomers served by server 2
    }
}

```

```

VipS1=0;           //number of VIP costumers served by server 1
RegS2=0;           //number of REG costumers served by server 2
now=0;
FreeS1=S1;
FreeS2=S2;
lambda1=1.3*row*(S1+1);
lambda2=0.4*row*(S1+1);

while(now<RunTime)
{
    double va=0,ra=0,vs=0,rs=0,vh=0,early_event=0;
    va=(rand()+1.)/(RAND_MAX+1.); //VIP arrivale
    VipArv=(-log(va)/lambda1,(
    ra=(rand()+1.)/(RAND_MAX+1.); //reg arrivale
    RegArv=(-log(ra)/lambda2,(

    if(FreeS1<S1)           //serv1 release VIP costomer
    {
        vs=(rand()+1.)/(RAND_MAX+1.);
        VipServ=(-log(vs)/(miy1*(S1-FreeS1)));
    }
    else VipServ=RunTime;
    serv2reg=S2-FreeS2-serv2vip;

    if(serv2reg>0)           //serv2 release reg costomer
    {
        rs=(rand()+1.)/(RAND_MAX+1.);
        RegServ=(-log(rs)/(miy2*serv2reg)));
    }
}

```

```

        else RegServ=RunTime;
if(serv2vip>0)                                //serv2 release VIP costomer
{
    vh=(rand()+1.)/(RAND_MAX+1);
    VipHelp=(-log(vh)/(miy3*serv2vip);
}
else VipHelp=RunTime;
early_event=Min(VipArv,RegArv,VipServ,RegServ,VipHelp);
now+=early_event;

if (early_event==VipArv)                      //arrival of Vip costomer
{
    total_VIP++;
    if (FreeS1>0)                            //there is a server1 avilable
    { FreeS1--;
      VipS1++;
    }
    else
    {
        if (FreeS2>0) //there is NO server1 avilable but there is a server2 avilable for VIP costomer
        { VipS2++;
          FreeS2--;
          serv2vip++; }
        else
        { VipQ++; //VIP stay in queue
        }
    }
}

```

```

if (early_event==RegArv)           //arrival of reg costomer
{
    total_reg++;
    if ((FreeS2>0)&&(VipQ==0))
    { RegS2++;
      FreeS2--; }
    Else
    RegQ++;
}

```

```

if (early_event==VipServ) //serv1 release VIP costomer
{
    FreeS1++;
    if (VipQ>0)
    {    VipQ--;
      FreeS1--;
      VipS1++; }
}

```

```

if (early_event==RegServ) //serv2 release reg costomer
{
    FreeS2++;
    if ((VipQ>0)&&(FreeS1==0))
    {
        VipQ--;
        FreeS2--;
        VipS2++;
        serv2vip++; }
}

```

```

        else if (RegQ>0)
        {
            RegQ--;
            FreeS2--;
            RegS2++; }
    }

    if (early_event==VipHelp) //serv2 release VIP costomer
    {
        FreeS2++;
        serv2vip--;
        if ((VipQ>0)&&(FreeS1==0))
        {
            VipQ--;
            FreeS2 --;
            VipS2++;
            serv2vip++; }
        else if (RegQ>0)
        {
            RegQ--;
            FreeS2--;
            RegS2++; }
    }

    //end while
}

VipS1_replication=VipS1_replication+VipS1;
VipS2_replication=VipS2_replication+VipS2;
RegS2_replication=RegS2_replication+RegS2;
total_VIP_replication=total_VIP_replication+total_VIP;
total_REG_replication=total_REG_replication+total_reg;
//end for of replication
}

```

```

Out1<<row<<'\\t'<<VipS1_replication/N_of_replication<<'\\t'<<VipS2_replication/N_of_replication<<'\\t'<
<RegS2_replication/N_of_replication<<'\\t'<<total_VIP_replication/N_of_replication<<'\\t'<<total_REG
_replication/N_of_replication<<endl;

```

```

//end for of row

```

```

}

```

```

}

```

```

return 0;

```

```

//end program

```

```

}

```

```

double Min(double VipArv, double RegArv, double VipServ, double RegServ, double VipHelp)

```

```

{

```

```

    double temp1=0;

```

```

    double temp2=0;

```

```

    double temp3=0;

```

```

    if (VipArv<RegArv) temp1=VipArv;

```

```

    else temp1=RegArv;

```

```

    if (VipServ<RegServ) temp2=VipServ;

```

```

    else temp2=RegServ;

```

```

    if (temp1<temp2)

```

```

    { if (VipHelp<temp1) return VipHelp;

```

```

        else return temp1; }

```

```

    else if (VipHelp<temp2)

```

```

        return VipHelp;

```

```

    else return temp2;

```

```

}

```

