การวิเคราะห์รูปแบบการกรองข้อความสั้นจากเนื้อหาร่วมกับการรับรองจากมนุษย์สำหรับการสื่อสารโทรศัพท์เคลื่อนที่

Analysis of Content Base & Human Intervention SMS Spam Filtering Model for Mobile Communication

อำนาจ ละมัยกลาง สุวิพล สิทธิชีวภาค

สาขาวิชาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

บทคัดย่อ

บริการข้อความสั้นหรือ Short Message Service (SMS) ได้รับความนิยมในการใช้งานเป็นอย่างมากในช่วงหลายปีที่ผ่านมาจึงทำให้จำนวนข้อความขยะเพิ่มขึ้นซึ่งส่งผลกระทบต่อประสิทธิภาพการทำงานของระบบศูนย์กลางบริการข้อความสั้นหรือ Short Message Service Center (SMSC) อย่างไรก็ดีเราสามารถควบคุมจำนวนข้อความขยะได้ด้วยระบบการกรองข้อความ บทความฉบับนี้จึงนำเสนอรูปแบบการกรองข้อความสั้นที่ผสมผสานระหว่างการตรวจสอบเนื้อหาและการรับรองจากมนุษย์ (CAPCHA) ของข้อความที่ไม่ได้ถูกจำแนกสถานะ หากไม่มีผลการตอบกลับแสดงว่าข้อความสั้นถูกส่งจากแหล่งที่สร้างข้อความขยะดังนั้นข้อความสั้นจึงไม่ถูกจัดส่งไปยังผู้รับปลายทาง เนื้อหาของบทความฉบับนี้จะอธิบายรูปแบบและกระบวนการทำงานที่สามารถจำแนกข้อความขยะ ผลการวิเคราะห์ทำงานของรูปแบบดังกล่าวพบว่ามีความน่าจะเป็นที่จะคัดแยกถูกต้อง 0.9354 ในขณะที่รูปแบบที่ทำการตรวจสอบเนื้อหาเพียงอย่างเดียวมีความน่าจะเป็นที่จะคัดแยกได้ถูกต้อง 0.9022 ดังนั้นรูปแบบที่นำเสนอนี้จะส่งผลให้ SMSC สามารถทำงานได้อย่างเต็มประสิทธิภาพมากยิ่งขึ้น

คำสำคัญ: ข้อความสั้น, ข้อความขยะ, ระบบศูนย์กลางบริการข้อความสั้น, การกรองข้อความ, การรับรองจากมนุษย์

Abstract

Short Message Service (SMS) has been the most popular means of mobile communication in recent years and hence the spam is an increasing threat to Short Message Service Center (SMSC) efficiency. The spam threat can be controlled through efficient and robust SMS filtering systems. In this paper we present new model that is a combination of content-base (CB) filtering and human intervention (CAPCHA). A message, that has been classified as uncertain by CB filtering, is further checked by sending a challenge to the message sender. An automated spam generator is unlikely to send back a correct response, in which case, the message is classified as spam and don't deliver to recipients. Based on this formulation, results show that our framework achieved a higher accuracy of 0.9354 comparing to those of content-based filtering at 0.9022 consequently, promoted efficiency of SMSC operation.

Keywords : SMS, Spam, SMSC, SMS Filtering, Human intervention

1.บทนำ

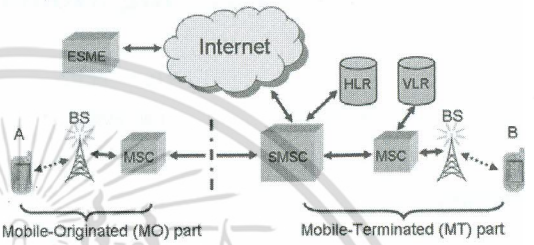
บริการส่งข้อความสั้นหรือ Short Message Service (SMS) รวมถึงบริการข้อความสื่อหรือ Multimedia Message Service (MMS) ได้รับความนิยมและเป็นที่แพร่หลายในการใช้บริการเป็นอย่างมากบนเครือข่ายโทรศัพท์เคลื่อนที่ (Mobile - Communication) เช่น เป็นเครื่องมือในการสื่อสารการตลาด, เป็นช่องทางในการตลาดแบบทางตรง, สร้างธุรกิจบริการเสริมหรือ Value Added Service (VAS)

เนื่องจากข้อความสั้นมีข้อดีหลายอย่าง เช่น เป็นสื่อที่มีประสิทธิภาพ, ค่าใช้จ่ายต่อข้อความมีแนวโน้มลดลง, สามารถกระตุ้นการรับรู้ได้ทันที ข้อความสั้นต่างๆ ที่ใช้ในบริการนี้อาจจะมีส่วนที่ถูกจัดกลุ่มเป็นข้อความขยะ (SMS Spam) ปะปนเข้ามาซึ่งจากการศึกษาในประเทศเกาหลีใต้และญี่ปุ่นนั้นพบว่ามีข้อความขยะสูงถึง 50% ของการใช้งาน ซึ่งส่งผลกระทบต่อประสิทธิ ภาพของศูนย์กลางการรับส่งข้อความหรือ Short Message Service Center (SMSC) ที่ต้องรับภาระการทำงานเกินความจำเป็น

Spam SMS คือข้อความสั้นที่ไม่ได้เรียกร้องให้ส่งหรือ Unsolicited Message ที่ก่อให้เกิดความรำคาญแก่ผู้ใช้และอาจสร้างปัญหาการล่อลวงให้เสียทรัพย์สินทางโทรศัพท์มือถือ เช่น ข้อความโฆษณาขายสินค้า-บริการ การหลอกล่อให้ผู้รับทำกิจกรรมบางประเภทที่เกิดความเสียหาย ข้อความหลอกลวง (Phishing) เป็นต้น หรือ Spammer อาจจะใช้ Robot Software เข้ามาช่วยในการส่ง Spam SMS ที่มีลักษณะการส่งครั้งละหลายข้อความและหลายปลายทางในครั้งเดียว ปัจจุบันได้มีมาตรการป้องกันต่างๆ เช่น การลงทะเบียนไม่ขอรับข้อความโฆษณาจากผู้ให้บริการ การใช้ Software กรองที่เครื่องโทรศัพท์ การใช้Software กรองที่ฝั่งเซิร์ฟเวอร์ เป็นต้น

การแก้ปัญหาด้วยซอฟท์แวร์การกรองข้อความขยะบนเครือข่ายโทรศัพท์เคลื่อนที่ที่ฝั่งเซิร์ฟเวอร์มีหลายหลายวิธี เช่น Bogofilter, DMC, LR, SVM [1] ซึ่งล้วนแต่มีการพัฒนาต่อเนื่องจากพื้นฐานการกรองอีเมลขยะ(E-Mail Spam Filtering) คือ การตรวจเนื้อหาและการจำแนก SMS นั้นว่าเป็นขยะหรือไม่ สำหรับงานวิจัยชิ้นนี้ได้นำวิธีการ Naive Bayesian [2] ที่มีการตรวจจับคำหรือวลีสำคัญซึ่งเป็นส่วนประกอบหนึ่ง SMS และเสนอรูปแบบการทำงานร่วมกันระหว่างการกรองข้อความสั้นจากเนื้อหา (Content-base) กับการรับรองของมนุษย์ที่ได้จากการถามตอบ (Challenge-response) ซึ่งเป็นส่วนหนึ่งของกลไกลอัตโนมัติที่มีวัตถุประสงค์เพื่อความปลอดภัยเนื่องจากการโจมตีจะใช้สิ่งที่เรียกว่า "บอตส์" (bots) ที่สร้างขึ้นจากคอมพิวเตอร์ แต่คอมพิวเตอร์ไม่สามารถแก้ปัญหาการทดสอบด้วย CAPTCHA ได้ ซึ่งมนุษย์เท่านั้นที่เพ่งดูกราฟฟิคและแกะตัวอักษรออกมาเพื่อพิมพ์ยืนยันรับรองเหล่า SMS นั้นๆ จากนั้นจึงนามน่าจะเป็นของการ SMS จาก CB ที่ได้มา ซึ่งอยู่ในรูปแบบความน่าจะเป็นไปวิเคราะห์แนวโน้มของการส่งผ่านข้อมูล (Traffic Path) ในกรณีต่างๆ ที่จะเกิดขึ้นได้ตามสมมติฐาน ภายใต้การจำลองการทำงานจากกลุ่มตัวอย่างของ SMS ที่ได้จากระบบโทรศัพท์เคลื่อนที่ที่ใช้งานจริงว่ารูปแบบการกรองข้อความที่นำเสนอนี้มีช่วยเพิ่มการกรองข้อความให้มีความถูกต้องมากขึ้น

2. ระบบการรับ-ส่งข้อความ



รูปที่ 1 โครงสร้างการรับ-ส่งข้อความ

รูปที่ 1 แสดงพื้นฐานของการรับ-ส่งข้อความในระบบโทรศัพท์เคลื่อนที่ประกอบด้วย 2 ส่วนสำคัญคือ

1.ผู้ส่ง (Mobile Originating : MO) ซึ่งรวมถึงเครื่องโทรศัพท์มือถือที่ใช้ส่ง, สถานีฐาน (Base Station : BS) และชุมสายโทรศัพท์มือถือ (Mobile Switching Center : MSC) ที่ทำหน้าที่ค้นหาเส้นทางและเชื่อมต่อสัญญาณต้นทาง-ปลายทาง

2. ผู้รับ (Mobile Terminating : MT) ซึ่งรวมสถานีฐาน (BS) และชุมสายโทรศัพท์มือถือ (MSC) ในส่วนปลายทาง นอกจากนี้ยังมีส่วนที่สำคัญคือ SMSC ที่ควบคุมระบบการรับส่ง SMS ทั้งหมดโดยจะได้รับข้อมูลตำแหน่งของผู้รับจาก Home Location Register (HLR) และ Visitor Location Register (VLR)

นอกจากนี้ระบบ SMS ยังสามารถรองรับการส่งอยู่อีกประเภทคือ External Short Message Entities (ESMEs) ที่เป็นการส่งจากหนึ่งผู้ส่งไปยังหลายผู้รับ (One-to-many message) ซึ่งได้ถูกนำมาใช้งานอย่างแพร่หลายในด้านการตลาดและบันเทิงเพราะคุ้มค่าจากการส่ง SMS ไปยังกลุ่มเป้าหมายได้ครั้งละจำนวนมาก เรียกการเชื่อมต่อแบบนี้ว่า Short Message Peer-to-Peer Protocal (SMPP Protocal) ซึ่งสามารถใช้งานผ่าน Internet ที่อาจจะถูก bot ใช้งานได้อย่างง่ายด้วยเช่นกัน

SMPP เป็น Protocal มาตรฐานในการรับ-ส่งข้อมูล SMS, MMS หรือ Push Message ภายในระบบ

โทรศัพท์เคลื่อนที่ซึ่งประกอบด้วย Protocal Description Unit (PDU) 2 ส่วนสำคัญดังตารางที่ 1 คือ ส่วนที่ 1 PDU Header ที่ใช้ระบุความยาวชนิดและลำดับของข้อความ ส่วนที่ 2 PDU Body ใช้บรรจุข้อมูลที่ต้องการส่ง เช่น เนื้อความ, ลิงค์สำหรับ Push Message เป็นต้น

ตารางที่ 1 รูปแบบของ SMPP PDU

SMPP PDU				
PDU Header (mandatory)				PDU Body (Optional)
Command length	Command id	Command status	Sequence number	PDU Body
4 octets	4 octets	4 octets	4 octets	Length*
4 octets	Command Length – 4			

Length*= (Command Length value - 16) octets

3. การกรองข้อความจากเนื้อหาร่วมกับการรับรองจาก มนุษย์ (Hybrid:การกรองแบบผสม)

รูปที่ 2 รูปแบบการกรองข้อความแบบผสม CB และ CAPCHA

3.1 การกรองข้อความจากเนื้อหา (CB Filtering)

การจำแนกประเภทข้อความด้วยวิธีการตรวจสอบเนื้อหาหรือ Content-Base (CB) นิยมใช้ Naive Bayesian ที่เป็นวิธีการจำแนกประเภทข้อมูลที่มีประสิทธิภาพเหมาะสมกับกรณีของเซตตัวอย่างมีจำนวนมากและมี Attribute ของตัวอย่างที่ไม่ขึ้นต่อกัน มีการนำไปประยุกต์ใช้งานในด้านการจำแนกประเภทข้อความ (Text Classification), การวินิจฉัย (Diagnois) ซึ่งพบว่าใช้งานได้ดี

กำหนดให้ตัวแปรสุ่ม y แทนกลุ่มข้อมูลข้อความ ที่มี Attribute ทั้งหมด n ตัว สามารถหาค่าความน่าจะเป็นของข้อความปกติแทนด้วย *ham* หรือ $Pr(c=ham|y)$ เขียนได้ดังนี้

$$p_r(c = ham \mid y) = \frac{p(ham)}{p(y)} \prod_i^n p(w_i = ham \mid y) \quad (1)$$

โดยที่ c คือความน่าจะเป็นของการจำแนกประเภทข้อความ

w_i คือลำดับของคำในข้อความ

ในทำนองเดียวกันความน่าจะเป็นของข้อความที่มีโอกาสเป็น SMS ขยะแทนด้วย *spam* หรือ $Pr(c=spam|y)$ เขียนได้ดังนี้

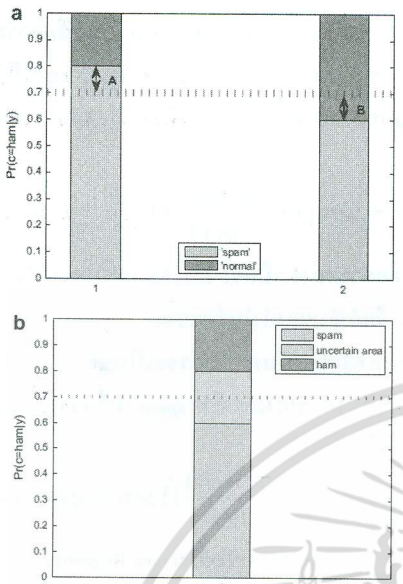
$$p_r(c = spam \mid y) = \frac{p(spam)}{p(y)} \prod_i^n p(w_i = spam \mid y) \quad (2)$$

3.2 การพิจารณาพื้นที่สีเทา (Uncertain Region)

ปกติระบบการกรองจาก CB สามารถจำแนกผลการทำงานได้เป็น 2 แบบคือ *ham* และ *spam* หากกำหนดความน่าจะเป็นของการกรองแทนด้วยการกระจายตัว $Pr(c=ham|y)$ แทนความน่าจะเป็นของข้อความที่อยู่ใน *ham region* โดย c และ y แทนตัวแปรสุ่มประเภทข้อความและข้อความตามลำดับ กำหนดอัตราส่วนที่ใช้ชี้วัดข้อความนั้นๆ ด้วย $O_{post} = Pr(c=ham|y)/Pr(c=spam|y)$ ถ้า $O_{post}>1$ แสดงว่าข้อความถูกจัดให้อยู่ใน *ham* และกรณีอื่นข้อความจะถูกจัดให้อยู่ใน *spam* เมื่อพิจารณาจุดอ้างอิง (Threshold-base) เพิ่มเพื่อใช้เป็นจุดแบ่งแยก จากกการที่ $Pr(c=ham|y)$ มีค่าเข้าใกล้ 1 แล้วข้อความน่าจะถูกจัดอยู่ใน *ham* และหากมีค่าเข้าใกล้ 0 จะถูกจัดอยู่ใน *spam* กำหนด $\bar{c} = f(y,h)$ เป็นค่าการกรองของ CB เมื่อ $\bar{c}$ คือเอาท์พุต และ h คือจุดอ้างอิง ดังนั้นเมื่อแทนค่าตัวแปรต่างๆ แล้วตัวกรองสามารถทำงานได้โดยใช้สมการดังนี้

$$\bar{c} = f(y,h) = \begin{cases} ham & if \quad \Pr(c=ham|y) \geq h \\ spam & if \quad \Pr(c=ham|y) < h \end{cases} \quad (3)$$

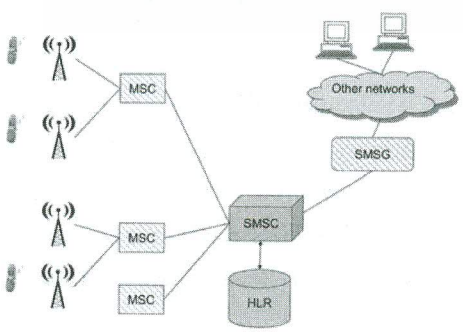
หากกำหนดจุดแบ่งแยกที่มีค่า $h=0.5$ อาจจะส่งผลทำให้เกิดปัญหาในการจำแนกประเภทข้อความได้ ดังนั้นเราจึงได้มีการนำขอบบนและขอบล่าง (Upper and Lower Boundaries) มาใช้ช่วยพิจารณาเพื่อแก้ไขจุดอ่อนของตัวกรอง CB ด้วยพื้นที่สีเทา (Uncertain Region) ดังแสดงในรูปที่ 3b



รูปที่ 3 (a) กรณีที่ 1 $h > \tilde{h}$ และกรณีที่ 2 $h < \tilde{h}$ โดยกำหนดค่าอ้างอิงจริง $\tilde{h}$ แทนด้วยเส้นประ (b) ปรับปรุงโดยการเพิ่มพื้นที่สีเทาและกำหนดค่าอ้างอิงจริง $\tilde{h}$ แทนด้วยเส้นประ

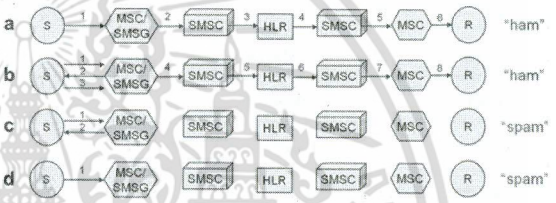
3.3 โปรโตคอลการถามตอบ:Challenge-response protocol

วิธีการจำแนกประเภทข้อความที่อยู่ในพื้นที่สีเทาว่าตกอยู่ในช่วงบวกและลบ (False Positive and False Negative) นิยมใช้กระบวนการ CAPTCHA [3] โดยตรวจสอบรูปแบบที่เข้ากันซึ่งผู้ใช้สามารถยืนยันได้ ส่วน SMS ขยะจากคอมพิวเตอร์หรือ bot อาจสร้าง SMS ขยะจำนวนมากได้แต่ไม่สามารถยืนยันตัวเองและตอบจากข้อความภาพเจือจางที่แสดงได้ ดังนั้นจึงกำหนดเมื่อมีการถามตอบที่ถูกต้องแสดงว่ามีความน่าจะเป็นสูงที่ SMSนั้นถูกส่งจากผู้ใช้ รูปแบบสื่อกลางของ CAPTCHA ที่สามารถปรับให้เหมาะสมกับการใช้งานได้ เช่น ภาพ เสียง หรือ ตัวอักษร เป็นต้น โดยจะเรียกวิธีการนี้ว่าการกรอง SMS แบบผสม (Hybrid)



รูปที่ 4 โครงสร้างการกรองข้อความแบบผสม (Hybrid)

โครงสร้างที่ออกแบบดังแสดงในรูปที่ 4 จำลองการทำงานโดยกำหนดผู้ส่ง (S), ผู้รับ (R), ศูนย์กลางบริการข้อความ (MSC หรือ SMSG) และส่วนประกอบอื่นๆ เช่น SMSC, HLR เป็นต้นโดยกำหนด $y^{h}_{c=Type}$ สำหรับ $type \in \{ham, spam\}$ แทนด้วยข้อความที่ถูกกรองซึ่งอยู่ในเทอมของ h ดังนั้นปริมาณข้อความรวมคือ $N_{FilteringOnly} = |y^{h}_{c=ham}| \times 6 + |y^{h}_{c=spam}| \times 1$ เมื่อ $|.|$ แทนจำนวนนับของ SMS ซึ่งปริมาณทั้งหมดในระบบ เนื่องจาก *ham* สามารถส่งผ่านไปยังโครงสร้างได้ 6 ส่วน คือ (S-MSC/SMSG-SMSC-HLR-SMSC-MSC-R) และ *spam* ส่งผ่านได้เพียง 1 ส่วนประกอบเท่านั้นคือ (S-MSC/SMSG)



รูปที่ 5 ความเป็นไปได้ของการส่ง SMS ในรูปแบบผสมทั้ง 4 กรณี

สำหรับรูปแบบผสม (Hybride) ที่ SMS ถูกแยกออกเป็น 3 ช่วงโดยใช้ 2 ค่า คือ h_1 และ h_2 จึงประมาณค่าพารามิเตอร์เพิ่มขึ้นมาอีก คือ N_{un} และ N_{us} ที่อยู่ในพื้นที่สีเทา ซึ่งสามารถแสดงความเป็นไปได้ของเส้นทางการส่งผ่านข้อมูลดังรูปที่ 5

กรณีที่ 1 (a) SMS ถูกจำแนกเป็น *ham* โดยให้ค่าความน่าจะเป็นสูงกว่าค่าอ้างอิงขอบบน จะถูกส่งไปยังผู้รับปลายทางผ่านอุปกรณ์ต่างๆ มีจำนวนการส่งผ่าน 6 ส่วน คือ S-MSC/SMSG-SMSC-HLR-SMSC-MSC-R

กรณีที่ 2 (b) SMS ถูกแยกอยู่ระหว่างขอบบนและขอบล่างในพื้นที่สีเทา เมื่อได้รับผลตอบที่ถูกต้องแล้ว SMS ถูกจำแนกเป็น *ham* มีการส่งผ่านทั้งหมด 8 ส่วน คือ S-MSC/SMSG-S-MSC/SMSG-SMSC-HLR-SMSC-MSC-R

กรณีที่ 3 (c) SMS ถูกแยกอยู่ระหว่างขอบบนและขอบล่างในพื้นที่สีเทา แต่ไม่ได้รับผลตอบที่ถูกต้อง แล้วข้อความถูกแยกเป็น *spam* มีจำนวนการส่งผ่าน 2 ส่วน คือ S-MSC/SMSG-S

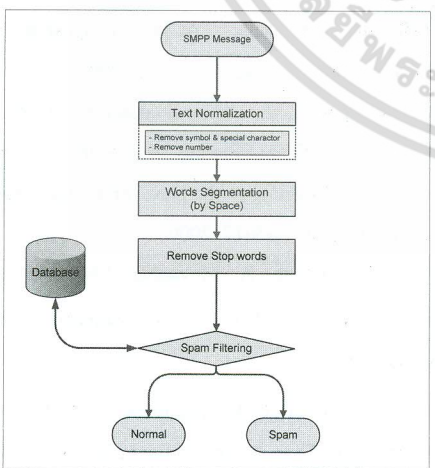
กรณีที่ 4 (d) ข้อความถูกจำแนกเป็นขยะ *spam* ไม่ถูกส่งต่อโดยคัดออกที่ศูนย์กลางบริการข้อความจึงมีจำนวนการส่งผ่าน 1 ส่วน คือ S-MSC/SMSG สามารถคำนวณปริมาณข้อมูลในโครงสร้างดังสมการ

$$
\begin{aligned}
N_n &= |\, y^{h_2}_{\bar{c}=ham} \,| \times 6 \\
N_{un} &= |\, y^{h_1}_{\bar{c}=ham} \cap y^{h_2}_{\bar{c}=spam} \cap y_{\bar{c}=spam} \,| \times (1-e_1) \times 8 \\
&\quad + |\, y^{h_1}_{\bar{c}=ham} \cap y^{h_2}_{\bar{c}=spam} \cap y_{\bar{c}=spam} \,| \times e_2 \times 8 \\
N_{us} &= |\, y^{h_1}_{\bar{c}=ham} \cap y^{h_2}_{\bar{c}=spam} \cap y_{\bar{c}=spam} \,| \times (1-e_2) \times 2 \\
&\quad + |\, y^{h_1}_{\bar{c}=ham} \cap y^{h_2}_{\bar{c}=spam} \cap y_{\bar{c}=ham} \,| \times e_1 \times 2 \\
N_{hybrid} &= N_n + N_{un} + N_{us} + N_s
\end{aligned}
\qquad (4)
$$

เมื่อ e_1 คือความน่าจะเป็นที่มีผลการตอบสนองกลับอย่างถูกต้อง (ส่งจากผู้ใช้) และ e_2 คือความน่าจะเป็นของขยะที่ถูกส่งจากคอมพิวเตอร์ที่อยู่ในพื้นที่สีเทา

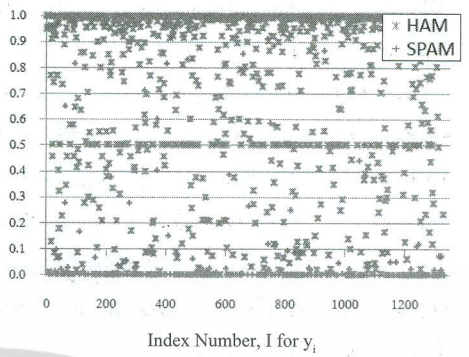
4. การทดสอบและการวิเคราะห์ผลการทดลอง

การทดลองการทำงานของตัวกรองเนื้อหา (CB Filtering) ในบทความฉบับนี้ได้จากการจำลองการทำงานจากโปรแกรมที่สร้างขึ้นมาและติดตั้งบนคอมพิวเตอร์แทนการติดตั้งที่ SMSC หรือ SMS Gateway โดยมีข้อมูลจำนวน 2 ชุดคือ ชุดข้อมูลฝึกสอน (Training Data:TD) ที่นำตัวอย่างข้อความขยะไปฝึกให้มีการเรียนรู้โดยชุดตัวอย่างคำที่เป็น SMS ขยะที่ได้มาจากผลการสำรวจของงานวิจัย [4] และชุดข้อมูลทดสอบชุดใหม่ (New Data:ND) ที่ผสมระหว่างภาษาไทยและอังกฤษซึ่งนำมาจากระบบบริการ CAT CDMA ของบริษัท กสท โทรคมนาคม จำกัด (มหาชน) ที่ให้บริการอยู่ในปัจจุบัน โดยมีขั้นตอนการทำงานดังรูปที่ 6



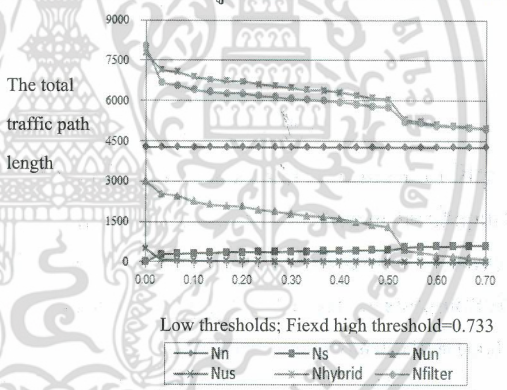
รูปที่ 6 ขั้นตอนการทำงานระบบ CB SMS Filtering

เมื่อระบบทำการคัดกรองเนื้อหาของ SMS แล้วและได้ค่า *Pr(c=ham|y)* หรือความน่าจะเป็นของ SMS ที่เป็นปกติและขยะซึ่งสามารถนำมาแสดงได้ดังรูปที่ 7

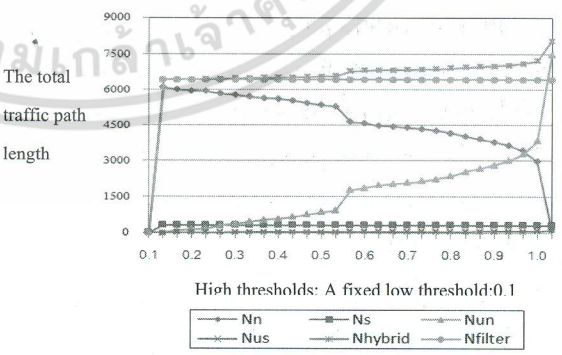


รูปที่ 7) การกระจายตัวของความน่าจะเป็น SMS

ผลจำลองการทำงาน CB Filtering พบว่ามี SMS ที่เป็นปกติ 82.2% และ SMS ขยะ 17.8% หากนำค่าความน่าจะเป็นของ SMS มาพิจารณาโดยใช้วิธีการรับรองรองจากมนุษย์มาเกี่ยวใช้ในการอ้างอิง(Threshold) จะถูกแบ่งออกเป็น 2 ช่วง ดังนั้นเราจึงนำค่า *Pr(c=ham|y)* ไปใช้เพื่อจำลองผลการวิเคราะห์ดังรูปแบบที่เสนอในบทความฉบับนี้



รูปที่ 8 Traffic Path เมื่อ h_2 คงที่ และ h_1 เปลี่ยนแปลงจาก 0 ถึง 0.733

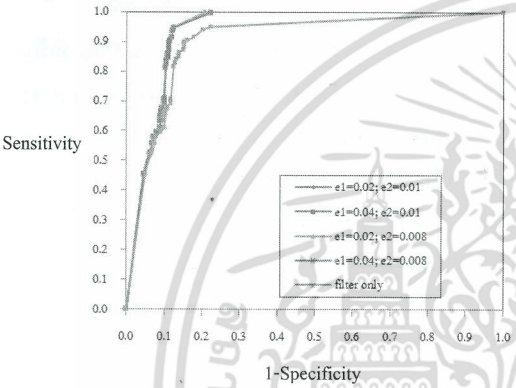


รูปที่ 9 Traffic Path เมื่อ h_1 คงที่ และ h_2 เปลี่ยนแปลงจาก 0.1 ถึง 1

รูปที่ 8 และ 9 แสดงผลของรวมของ Traffic path ที่มีการกำหนดค่า h_1, h_2 จากกราฟแสดงให้เห็นว่าการ Traffic usage ลดลงเมื่อค่า h_1 มีค่าเพิ่มมากขึ้นซึ่งก็สอดคล้องกับความเป็นจริงที่เกิดขึ้นคือ h_1 ที่มากขึ้นนั้น SMS จะถูกแยก

เป็น *spam* และส่งผลให้ถูกนำไปสู่รับรองจึงทำให้จำนวน Traffic path ลดลง

อีกตัวชี้วัดหนึ่งเพื่อแบ่งแยก SMS คือ Receiver Operating Characteristic (ROC) Curve ที่ได้นำค่าซึ่งตกอยู่ในพื้นที่สีเทามาเปรียบเทียบ คือ True Positive (TP), True Negative (TN), False Positive (FP) และ False Negative (FN) ที่มีค่าตั้งแต่ 0 ถึง 1 โดยเปรียบเทียบระหว่าง CB Filtering ที่มี h เป็นค่าอ้างอิงเพียงค่าเดียวและแบบผสมที่มีค่า h_1 และ h_2 มาช่วยคัดแยก SMS



รูปที่ 10 ROC Curve

รูปที่ 10 กราฟแสดง ROC Curve ที่บอกถึงแนวโน้มความถูกต้องของการคัดแยก SMS โดยที่แกน x แทนค่าของ 1-*Specificity* และแกน y แทนค่าของ *Sensitivity* ที่เปรียบเทียบผลจากการทำงานของระบบ CB Filtering และ Hybride ซึ่งประมาณค่าจากการสมการ

$$Specificity = \frac{TN}{FP+TN} \quad and \quad sensitivity = \frac{TP}{TP+FN} \qquad (5)$$

พบว่าการกรองแบบผสมนั้นมีค่าความถูกต้องที่สูงกว่าแบบ CB Filtering และเมื่อค่า e_1 และ e_2 มีค่าน้อยลงส่งผลให้เส้น ROC มีค่าเพิ่มขึ้น ซึ่งเมื่อพิจารณาพื้นที่ใต้ ROC Curve หรือ Area Under the ROC Curve (AUC) ที่บ่งบอกความน่าจะเป็นของการคัดแยก SMS ที่จะมีค่าเป็นปกติมากกว่า SMS ขยะ

ตารางที่ 2 การเปรียบเทียบค่า AUC

Method	Ratio (e_1)	Ratio (e_2)	AUC
Filter only	-	-	0.9022
Hybrid	0.02	0.008	0.9354
Hybrid	0.04	0.008	0.9347
Hybrid	0.02	0.01	0.9351
Hybrid	0.04	0.01	0.9344

จากตารางที่จะเห็นได้ว่า AUC ของรูปแบบผสมนั้นมีค่า 0.9354 ซึ่งมากกว่าแบบ CB Filtering ที่ 0.9022 โดยเป็นการเพิ่มความน่าจะเป็นที่ระบบจะกรอง SMS ปกติได้ถูกต้องมากขึ้น ค่าของ e_1 และ e_2 ที่เพิ่มขึ้นส่งผลให้ AUC มีค่าลดลง

5. สรุปผลการทดลอง

บทความนี้เสนอวิธีการกรอง SMS แบบผสมระหว่าง CB Filtering และการรับรองจากมนุษย์ (CHAPCHA) จากรูปแบบภาพเจือจางที่ Bot ไม่สามารถตอบได้ โดยวิเคราะห์ผลการจำลองการทำงานจากจำนวนการส่งผ่านข้อมูลแล้วนำมาหาค่าความน่าจะเป็นของ SMS ปกติอย่างเพื่อแสดงให้เห็นว่าการกรอง SMS ที่นำเสนอนี้ช่วยให้สามารถกรอง SMS ได้ถูกต้องมากกว่าการนำ SMS มาการกรองแบบ CB Filtering เพียงอย่างเดียว ซึ่งเป็นการลดภาระการทำงานในส่วนของข้อมูลขยะที่ไม่จำเป็นให้ทำงานได้อย่างเต็มประสิทธิภาพมากขึ้น ทั้งนี้ค่าต่างๆ ที่จะนำมาใช้งานจะต้องคำนึงถึงค่าอ้างอิง h , h_1 , h_2 , e_1 , e_2 ที่เหมาะสมเนื่องจากจะส่งผลกระทบต่อจำนวนของ SMS ที่ถูกนำมาเข้าสู่กระบวนการรับรองจากมนุษย์เพิ่มเติมมากเกินความจำเป็นซึ่งอุปกรณ์ก็จะต้องทำงานได้อย่างรวดเร็วและถูกต้อง จึงจะสามารถช่วยลดภาระการทำงานของ SMSC ได้ตามวัตถุประสงค์

6. เอกสารอ้างอิง

[1] G. Hidalgo, J. Maria, E. Sanz, Gacia, "Content base SMS spam filtering," ACM, Symposium on Document engineering, pp.114-122, 2006.

[2] Androutsopoulos I., "An Evaluation of Naive Bayesian Anti-Spam Filtering, Proc of the workshop on Machine Learning in the New Information Age," Bacelona, Spain, pp.9-17, 2000.

[3] S. Shirali-Shahreza, A. Movaghar, "An anti-spam using CAPTCHA," IEEE Trans. Computer Society, pp. 318-321, 2008.

[4] N. Boonnitiprasert, C. Khemmapatapan, "SMS Filtering for Thai & EnglishLanguage on Mobile Phone Network," NCCIT. The 5th National Conference on Computing and Information Technology, Bangkok, pp.436-442, 2009.