

ลายพิมพ์เสียง : เทคโนโลยีเพื่อการค้นหาเพลง

อัครพร แพงดี และ พรฤดี เนติโสภาคกุล

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

Email : agumi.p@gmail.com

บทคัดย่อ

ในการสืบค้นข้อมูลในอดีตส่วนใหญ่แล้วเราจะนึกถึงการป้อนคำสำคัญ ที่เราต้องการค้นหาเข้าไปในโปรแกรมค้นหา หลังจากนั้นโปรแกรมค้นหาจะแสดงรายการของผลลัพธ์ที่ใกล้เคียงกับคำสำคัญมากที่สุดขึ้นมาให้เราเลือกใช้งาน ในปัจจุบันการสืบค้นข้อมูลไม่ได้จำกัดอยู่แค่การใช้ตัวหนังสือในการค้นหาเท่านั้น โดยเฉพาะการค้นหาข้อมูลบนอินเทอร์เน็ตสามารถใช้เสียงมาแทนคำสำคัญในการค้นหาข้อมูล โดยโปรแกรมค้นหาสามารถฟังเสียงและนำไปใช้ค้นหารายละเอียดได้ไม่ว่าเสียงนั้นจะเป็นเสียงเพลง หรือเสียงพูดของมนุษย์ ความมหัศจรรย์นี้ได้มาจาก เทคโนโลยีที่มีชื่อว่า ‘Audio Fingerprinting’ หรือ ‘ลายพิมพ์เสียง’ ซึ่งสามารถนำมาใช้ประโยชน์ในงานที่เกี่ยวข้องกับการระบุข้อมูลหรือลิขสิทธิ์ของมัลติมีเดียต่างๆ ทำให้การตรวจสอบการละเมิดลิขสิทธิ์ของเพลงบนอินเทอร์เน็ต หรือการเปิดเพลงของคลื่นวิทยุได้สะดวกมากยิ่งขึ้น โดยผู้ให้บริการสืบค้นข้อมูลเสียงเพลงบนอินเทอร์เน็ตที่นำเทคโนโลยีนี้มาใช้ในปัจจุบันเริ่มมีการขึ้น เช่น โปรแกรมประยุกต์ของบริษัท Shazam และ SoundHound ที่เน้นการให้บริการสืบค้นเสียงเพลงและเชื่อมต่อไปยังบริการซื้อเพลงบนมือถือได้ และ ผู้ให้บริการทางด้านการตรวจสอบลิขสิทธิ์ของเสียงเพลงผ่านระบบเครือข่ายต่างๆ ยกตัวอย่างเช่น Audible Magic , Music Trace และ Broadcast Data System เป็นต้น

คำสำคัญ – Audio Fingerprinting; ลายพิมพ์เสียง; Shazam; SoundHound; Audible Magic; Music Trace; Broadcast Data System

1. บทนำ

เทคโนโลยีการสืบค้นข้อมูล โดยเฉพาะข้อมูลบนอินเทอร์เน็ต ในปัจจุบันส่วนใหญ่ทำได้โดยป้อนคำสำคัญ (คีย์เวิร์ด) ให้กับ โปรแกรมค้นหา (Search Engine) จากนั้น โปรแกรมค้นหา ก็จะแสดงรายการของผลลัพธ์ที่ผู้ใช้นั้นต้องการขึ้นมา คงไม่ใช่เรื่องยากหากคำสำคัญที่ต้องการหาอยู่นั้นอยู่ในรูปของตัวหนังสือ แต่หากคำสำคัญที่เราต้องการค้นหาเป็นเสียง เทคโนโลยีสืบค้นข้อมูลเมื่อ 10 ปีก่อนคงไม่มีทางจะหาผลลัพธ์ที่ต้องการได้ หากแต่ในปัจจุบันมีการค้นหาข้อมูลด้วยเสียงมิใช่เรื่องที่เป็นไปไม่ได้ อีกต่อไปแล้ว เนื่องจากมีเทคโนโลยีที่มีชื่อว่า ‘Audio

Fingerprinting’ เข้ามาช่วยให้การสืบค้นข้อมูลให้เป็นไปอย่างสมบูรณ์ยิ่งขึ้น

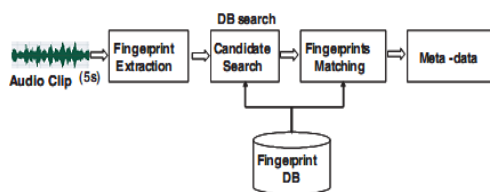
2. Audio Fingerprinting คืออะไร?

Audio Fingerprinting คือลายพิมพ์ของข้อมูลมัลติมีเดียสามารถระบุได้ถึงเนื้อหาของมัลติมีเดียได้เช่นเดียวกับลายนิ้วมือของมนุษย์ที่ใช้ระบุถึงตัวบุคคล แต่ในที่นี้ใช้ระบุคลิปเสียง เสียงของเครื่องดนตรีต่างๆ รวมไปถึงเสียงของมนุษย์ [6] ซึ่งเสียงดังกล่าวถูกแปลงจากสัญญาณเสียงแอนะล็อกเป็นสัญญาณดิจิทัล โดยใช้ทฤษฎีทางคณิตศาสตร์ที่ซับซ้อนและการรวบรวมฐานข้อมูลขนาดใหญ่ ที่เก็บข้อมูลของลายพิมพ์เสียงต้นฉบับ ที่ผ่านการบีบอัดข้อมูล

อย่างมีประสิทธิภาพ เพื่อนำไปใช้คำนวณเทียบเคียงกับ ตัวอย่างเสียงที่ต้องการได้อย่างถูกต้อง[7]

2.1. System Overview

การทำให้ลายพิมพ์เสียงสามารถนำไปใช้งานได้อย่างมีประสิทธิภาพนั้น จะต้องประกอบไปด้วยขั้นตอนการทำงานที่สำคัญดังนี้ 1) Fingerprinting Extraction 2) Candidate Search และ Fingerprints Matching 3) Meta – data



รูปที่ 1. Audio Fingerprinting System

จากรูป 1 ระบบ Audio Fingerprinting [8] โดยทั่วไปจำแนกได้เป็นขั้นตอน ดังนี้

2.1.1. Fingerprinting Extraction หรือ Feature Extraction

ทำหน้าที่รับเสียงเพลงเป้าหมายมาจากภายนอก เพื่อหา ลายพิมพ์เสียง (feature) หรือ การแปลงเสียงที่ได้รับเข้ามาให้อยู่ในรูปที่คอมพิวเตอร์สามารถประมวลผลได้

วิธีการทั่วไปคือการแยกเสียงเป็น เฟรม จากนั้นคำนวณคุณสมบัติของแต่ละเฟรม จากทฤษฎีทางคณิตศาสตร์ได้แก่ สัมประสิทธิ์ฟูเรียร์ (Fourier coefficients) , Mel-Frequency Cepstral Coefficients (MFCC), การประมาณพั่นระเชิงเส้น (LPC) เป็นต้น [7] ถ้ามองในทางปฏิบัติแล้วเนื้อหาใน มัลติมีเดียจะถูกบีบอัดและเพิ่มประสิทธิภาพของเสียง อันเนื่องมาจากการแปลงสัญญาณจากแอนะล็อกเป็น ดิจิทัล ดังนั้นระบบลายพิมพ์เสียงจะต้องสามารถ อนุญาตให้มีการ การปรับเปลี่ยนของเสียง เช่น ปรับลด

สัญญาณรบกวน หรือ การขยายสัญญาณ เป็นต้น โดยทั่วไปแล้วฟังก์ชันลายพิมพ์ต้องมีคุณสมบัติดังต่อไปนี้ [6]

- Robustness : ลายพิมพ์ของคลิปเสียงไม่สมบูรณ์ควรใกล้เคียงกับ ลายพิมพ์ของคลิปเสียงต้นฉบับ
- Pair wise independence : สองคลิปเสียงที่แตกต่างกัน จะต้องมีการรับรู้ลายพิมพ์ที่แตกต่างกัน
- Database search efficiency : โครงสร้างของลายพิมพ์เสียงจะต้องมีฐานข้อมูลที่เอื้อต่อการค้นหาที่รวดเร็ว

2.1.2. Candidate Search และ Fingerprints Matching

เป็นส่วนที่นำลายพิมพ์เสียงจากตัวอย่าง มาเปรียบเทียบกับข้อมูลของลายพิมพ์เสียง ในฐานข้อมูลที่มีอยู่ เพื่อหา เพลงที่มีความใกล้เคียงกับตัวอย่างที่สุดในช่วงเวลาเดียวกัน

ขั้นตอนการค้นหาลายพิมพ์เริ่มจาก สร้างตารางแฮชของแคตตาล็อกลายพิมพ์ของตัวอย่าง โดยใช้ความถี่สูงสุดเป็น คีย์แรกในการค้นหา แฮชที่มีความถี่ที่ใกล้เคียงกัน [10]

จากตารางที่ 1 และ 2 จะเห็นได้ว่าในช่วงความถี่ที่ 823.44 และ 1892.31 ในตาราง hash ของเพลงตัวอย่าง ตรงกับความถี่ในตาราง hash ของฐานข้อมูล ที่วินาทีที่ 34.678 และ 34.945 ของ “Song B” by Artist 2 [10]

ตารางที่ 1. แสดง Hash Table ตัวอย่างที่ 10 วินาที

Frequency in Hz	Time in seconds
823.44	1.054
1892.31	1.321
712.84	1.703
.....	
819.71	9.943

ตารางที่ 2. แสดง Hash Table ของฐานข้อมูล

Frequency in Hz	Time in seconds, song information
823.43	53.352, "Song A" by Artist 1
823.44	34.678, "Song B" by Artist 2
823.45	108.65, "Song C" by Artist 3
.....	
1892.31	34.945, "Song B" by Artist 2

ฐานข้อมูลของเพลง (Fingerprint DB) เป็นฐานข้อมูลซึ่งใช้เก็บรายละเอียดของเพลง พร้อมทั้งลายพิมพ์เสียงที่จะใช้ในการเปรียบเทียบเพื่อค้นหาโดยลายพิมพ์เสียงที่เก็บภายในฐานข้อมูลนั้นต้องมีความสมบูรณ์ปราศจากสัญญาณรบกวน เพื่อความแม่นยำในการค้นหาเพลงจากตัวอย่างเพลงที่มีสัญญาณรบกวนสูง

2.1.3. Meta – data หรือ Output

ผลลัพธ์จากการค้นหาประกอบด้วย ชื่อเพลง ชื่อศิลปิน หรือรายละเอียดอื่น ๆ ที่อยู่ในฐานข้อมูล

2.2. ปัจจัยที่ส่งผลต่อประสิทธิภาพของ Audio Fingerprinting System [1]

- 1) Noise : เมื่อรับเสียงเพลงเป้าหมายในสภาพแวดล้อมทั่วไปที่มีเสียงดังหรือเสียงรบกวนมากๆ เกิดสัญญาณเสียงที่ไม่พึงประสงค์เข้ามาพร้อมกับสัญญาณเสียงเป้าหมาย หรืออาจกลบเสียงที่เราต้องการ ทำให้สัญญาณที่ส่งเข้าไปในระบบเปลี่ยนแปลงไปจากเดิม ระบบต้องมีขั้นตอนวิธีการที่จะจัดการกับเสียงรบกวนเหล่านี้อย่างเหมาะสม
- 2) Distortion : ระบบจะต้องสามารถจัดการกับการบิดเบือนที่เกิดขึ้นจากแหล่งต่าง ๆ รวมทั้งการเล่นที่ไม่สมบูรณ์เกิดจากอุปกรณ์ที่ใช้เก็บตัวอย่างเอง เนื่องจากการเก็บตัวอย่างผ่านอุปกรณ์โทรศัพท์นั้น มีการลดทอนความถี่ตั้งแต่ 300Hz - 3, 400Hz ทำให้สัญญาณตัวอย่างเสียงที่ส่งเข้าไปในระบบผิดเพี้ยนไป

ดังนั้นหากรับสัญญาณเสียงตัวอย่างจากอุปกรณ์มือถือที่มีประสิทธิภาพ ก็จะส่งผลให้ได้สัญญาณที่มีประสิทธิภาพด้วยเช่นกัน

- 3) Database management : ระบบสามารถทำงานได้ดีเมื่อจำนวนของลายพิมพ์เสียงในฐานข้อมูลมีมากขึ้น โดยที่ประสิทธิภาพการทำงานในการค้นหาเสียงเพลงยังคงทำได้อย่างรวดเร็ว และถูกต้องเสมอ

2.3. ประโยชน์ของ Audio Fingerprint System

ประการแรกลายพิมพ์เสียงช่วยให้สามารถหา meta-data ของเพลงใน library ที่สอดคล้องกัน ซึ่งช่วยให้องค์กรได้ใช้งานง่ายขึ้น ตัวอย่างเช่นอัลบั้ม ศิลปิน เนื้อเพลง ลิงก์สำหรับดาวน์โหลดเพลงหรือข้อมูลอื่นๆ ที่ถูกเก็บอยู่ใน meta-data

ประการที่สองสามารถช่วยตรวจสอบการดาวน์โหลดหรืออัปโหลดเพลงให้เป็นไปอย่างถูกต้องตามลิขสิทธิ์ สิ่งสำคัญที่สุดคือ ขั้นตอนวิธีการมีประสิทธิภาพพอที่จะให้ทำงานกับอุปกรณ์พกพาขนาดเล็ก [7]

ในอนาคตเทคโนโลยี Audio printing นี้อาจจะเข้ามามีส่วนช่วยอย่างมากในการเฝ้าติดตามวิเคราะห์ข้อมูลเกี่ยวกับข้อมูลในวงการเพลงและโฆษณาบนสถานีวิทยุและโทรทัศน์ ไม่เพียงแต่ผลดีทางด้านการตรวจสอบ meta-data ของเพลงเพื่อระบุลิขสิทธิ์การดาวน์โหลดเพลงเท่านั้น อาจมีการพัฒนาระบบเพื่อช่วยวิเคราะห์ข้อมูลเพื่อการวิจัยตลาด และเพื่อประกอบการตัดสินใจในการวางแผนการลงทุนทั้งในอุตสาหกรรมเพลงและสื่อโฆษณาได้อีกด้วย

3. ตัวอย่างโปรแกรมประยุกต์และบริการที่นำเทคโนโลยี Audio Fingerprinting ไปใช้

เทคโนโลยี Audio Fingerprinting มีการนำไปใช้งานมาก ทั้งในธุรกิจภาคเอกชน และโดเมนสาธารณะ เช่น โปรแกรมประยุกต์ของบริษัท Shazam (<http://www.shazam.com/>) และ SoundHound

(<http://www.soundhound.com/>) โดยทั้งสองบริษัทได้นำเสนอแอปพลิเคชันที่ใช้เทคโนโลยี Audio Fingerprinting ผ่านทางโทรศัพท์มือถือซึ่งทำให้ผู้ใช้สามารถใช้งานแอปพลิเคชันสำหรับดูรายละเอียดหรือเลือกซื้อเพลงที่ต้องการในสถานที่สาธารณะได้

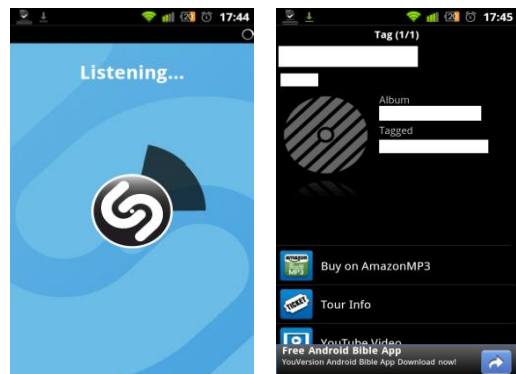
ในหลายประเทศสถานีวิทยจะต้องจ่ายค่าลิขสิทธิ์สำหรับเพลงที่จะออกอากาศของแต่ละสถานี เจ้าของลิขสิทธิ์จะสามารถตรวจสอบการส่งสัญญาณวิทยุ ของแต่ละสถานีวิทยได้ว่าเพลงที่ออกอากาศอยู่นั้นได้จ่ายค่าลิขสิทธิ์อย่างถูกต้องแล้วหรือยัง หรือ สังเกตการณ์เกี่ยวกับการเปิดเพลงหรือโฆษณา เพื่อตรวจสอบว่ามีการออกอากาศโฆษณาหรือเปิดเพลงตามที่ตกลงไว้หรือไม่ โดยนำเทคโนโลยี Audio Fingerprinting ไปใช้ในการเทียบเคียงเสียง และระบุเพลงที่ถูกเปิดเล่นในเวลาเดียวกันจากสถานีวิทยต่างๆ แล้วจึงรายงานผลการสำรวจเกี่ยวกับการเปิดเล่นเพลง และโฆษณา ให้กับผู้โฆษณา ตัวอย่างของผู้ให้บริการเชิงพาณิชย์ของบริการนี้คือ Audible Magic (www.audiblemagic.com), Music Trace (<http://www.musictrace.de>) และ Broadcast Data System (www.bdsnline.com) เป็นต้น[4]

3.1. ตัวอย่างแอปพลิเคชัน Shazam

Shazam คือ บริการที่สามารถระบุเพลง จำแนกเพลงจากการฟังเพลงในระยะเวลา 10 วินาที จากนั้นจะส่งเพลงส่วนที่ฟังนั้นไปวิเคราะห์บนเซิร์ฟเวอร์ที่มีประสิทธิภาพและประมวลผล ให้ข้อมูลชื่อเพลง ชื่อศิลปิน อัลบั้ม เนื้อเพลง และสามารถสั่งซื้อเพลงผ่านทางบริการนี้ได้ ผู้ใช้สามารถเรียกใช้บริการผ่านทาง แอปพลิเคชัน ของ Shazam บนโทรศัพท์เคลื่อนที่ได้ทุกที่ [1],[2] ที่ต้องการ ซึ่งในปัจจุบันสามารถรันได้หลายแพลตฟอร์ม ทั้ง iOS, Android, BlackBerry, Windows Mobile, Windows Phone 7, และ Symbian มีให้เลือกใช้ทั้งรุ่นฟรี และ เสียค่าใช้จ่าย

Shazam Entertainment, Ltd เริ่มต้นในปี ค.ศ. 2000 มีความคิดในการให้บริการที่สามารถเชื่อมต่อผู้คน

กับเสียงเพลงโดยใช้ไมโครโฟนของโทรศัพท์เคลื่อนที่รับตัวอย่างของเสียงเพลงภายใต้สภาพแวดล้อมที่อาจมีเสียงรบกวน แปลงสัญญาณและการบีบอัดจากเครือข่าย ก่อนจะส่งสัญญาณเสียงดังกล่าวมาที่เซิร์ฟเวอร์ [3] เพื่อหาข้อมูลของเสียงนั้นและส่งผลลัพธ์กลับไปยังหน้าจอของผู้ใช้

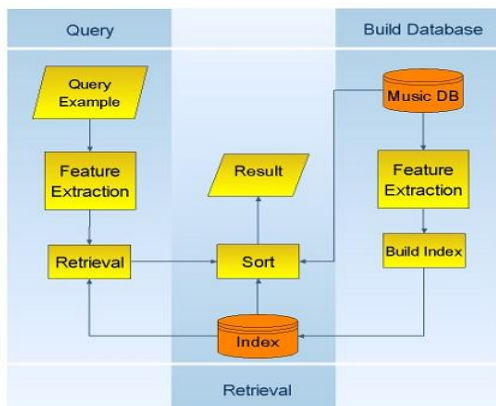


รูปที่ 2. ตัวอย่าง แอปพลิเคชัน Shazam

เทคโนโลยีเบื้องหลัง Shazam
ขั้นตอนของ Shazam

- 1) Shazam สร้างและเก็บข้อมูลเกี่ยวกับเพลง รวมทั้งลายพิมพ์เสียง (Fingerprint) ของเพลงต้นฉบับไว้ในฐานข้อมูล
- 2) ผู้ใช้กดปุ่ม 'Tag' เพลงบนแอปพลิเคชัน เพื่อส่งตัวอย่างของเพลงที่ต้องการไปยังฐานข้อมูลของ Shazam ผ่านทางไมโครโฟน
- 3) Server ของ Shazam ค้นหาเพลงที่มีลายพิมพ์เสียงตรงกับเพลงตัวอย่างของผู้ใช้ในฐานข้อมูลและส่งผลลัพธ์ที่ได้ออกไปแสดงผลที่หน้าจอของผู้ใช้

จากรูปที่ 2 ระบบของ Shazam จะแบ่งเป็น 2 ส่วนสำคัญ คือ Build Database process และ Retrieval process. Build Database process ทำหน้าที่แปลง



รูปที่ 3. โครงสร้าง ของ Shazam System

เพลงให้อยู่ใน รูปแบบดิจิทัลเพื่อเก็บไว้ในฐานข้อมูล ส่วน Query process จะส่งสัญญาณเสียงตัวอย่างที่ถูกแปลง ให้อยู่ในรูปแบบเดียวกับฐานข้อมูลเข้ามาที่ Retrieval process เพื่อเปรียบเทียบกับสัญญาณเสียงที่มีอยู่ใน ฐานข้อมูล [5]

3.2. ตัวอย่างการให้บริการตรวจสอบลิขสิทธิ์จาก Audible Magic

Audible Magic เป็นผู้นำในด้านการให้บริการทางด้าน ลิขสิทธิ์ ที่มีลูกค้ากว่า 150 ราย โดยมีฐานข้อมูลของ Audible Magic ประกอบไปด้วยเพลง ภาพยนตร์ และ รายการโทรทัศน์ ที่ถูกต้องตามลิขสิทธิ์ครอบคลุมกว่า 12 ล้านเพลง ที่ค่ายศิลปินและสตูดิโอภาพยนตร์โทรทัศน์จาก ทั่วโลกมาลงทะเบียนกับ Audible Magic เพื่อให้ มัลติมีเดียของพวกเขาได้รับการป้องกันจากการละเมิด ลิขสิทธิ์[9]

บริการของทาง Audible Magic นี้จะใช้ข้อมูลของทั้ง ลายพิมพ์เสียง และลายพิมพ์ภาพเคลื่อนไหว มาหาความ แตกต่างของสื่ออื่นๆ กับลายพิมพ์เสียงและลายพิมพ์ ภาพเคลื่อนไหว ที่มีในฐานข้อมูล เพื่อระบุเนื้อหาของเสียง และวิดีโอที่ต้องการ



รูปที่ 4. ตัวอย่างเว็บไซต์ Audible Magic

เทคโนโลยีที่นำมาใช้ในบริการของ Audible magic ประกอบด้วย 2 เทคโนโลยี คือ

- 1) Audio Fingerprinting เทคโนโลยีลายพิมพ์เสียงใช้ ในการระบุรายละเอียดของเสียงเพลงหรือโฆษณา ผ่านทางการรับรู้ด้วยเสียงเช่นเดียวกับมนุษย์และ เปรียบเทียบลักษณะเฉพาะของแต่ละสัญญาณเสียงที่ อยู่ในฐานข้อมูล ซึ่งขั้นตอนการนำข้อมูลลายพิมพ์เสียง เข้าไปเก็บในฐานข้อมูล [11]
- 2) Audio Watermarking เทคโนโลยีลายน้ำเสียงใช้ในการ ช่วยฝังข้อมูลบางอย่างเพิ่มเติมในแทร็กเสียง ทำ ให้สามารถตรวจจับการสำเนาข้อมูลเสียงที่ผิด กฎหมายได้ง่ายขึ้น โดยหากพบลายน้ำเสียงที่ไม่ ถูกต้องจะไม่สามารถเข้าถึงข้อมูลของเพลงหรือวิดีโอ นั้นๆได้ เทคโนโลยีลายน้ำเสียงจึงเป็นเทคโนโลยีที่มี ความน่าเชื่อถืออย่างมาก นอกจากการป้องกันไม่ให้ เข้าถึงข้อมูลได้แล้ว ลายน้ำเสียงยังมีบทบาทสำคัญ ในการบ่งชี้ปริมาณการเข้าถึงข้อมูลในแต่ละช่วงเวลา ได้อีกด้วย อย่างไรก็ตามหากเสียงที่นำมาตรวจสอบ ผ่านการแปลงสัญญาณซ้ำๆ จนทำให้ไม่สามารถสกัด ลายน้ำเสียงได้ จะไม่สามารถตรวจสอบ สัญญาณเสียงด้วยลายน้ำเสียงได้ [12]

Social networking อย่างเช่น My Space , Facebook ได้นำโปรแกรมที่ออกแบบโดยบริษัท Audible Magic ไปใช้เพื่อตรวจสอบการโพสวิดีโอที่ ละเมิดลิขสิทธิ์ไม่ให้ถูกเผยแพร่ทางอินเทอร์เน็ต

4. บทสรุป

ถึงแม้ Audio fingerprint system จะไม่ใช่เทคโนโลยีใหม่แต่ก็เป็นอีกเทคโนโลยีที่น่าสนใจ เนื่องจากสามารถนำไปใช้ประโยชน์ได้หลายด้าน ทั้งในด้านตรวจสอบลิขสิทธิ์เกี่ยวกับเพลง การเก็บสถิติในการโฆษณา หรือแม้แต่การระบุรายละเอียดเพลงจากการฟังเพลงนั้นๆ ตามสถานที่ต่างๆ ในอนาคตหากมีการใช้เทคโนโลยีนี้อย่างแพร่หลายมากขึ้น อาจช่วยให้การละเมิดลิขสิทธิ์เพลงบนอินเทอร์เน็ตก็จะทำได้ยากขึ้น ซึ่งเป็นผลดีอย่างยิ่งกับวงการเพลงและโฆษณานบนอินเทอร์เน็ต

เอกสารอ้างอิง

- [1] Avery Li-Chan Wang, 2006. The Shazam Music Recognition Service. COMMUNICATIONS OF THE ACM , Vol 49, No 8.
- [2] Roy Want, 2010, iPhone: Smarter Than the Average Phone. Pervasive Computing, IEEE , Vol 9
- [3] Avery Li-Chan Wang. 2003. An Industrial-Strength Audio Search Algorithm. Retrieved Jan 17, 2011, from <http://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf>
- [4] Cano P. et al. 2005. Audio Fingerprinting Concepts and Applications. Retrieved Apr 11, 2012, from <http://mtg.upf.edu/files/publications/0e9cd9-Springer05-pcano.pdf>
- [5] Tang Jie and Liu Gang , Guo Jun. 2009. Improved Algorithms of Music Information Retrieval based on Audio Fingerprint, IEEE Intelligent Information Technology.
- [6] Jin S. Seo et al. 2006. Audio Fingerprinting Based on Normalized Spectral Subband Moments. IEEE SIGNAL PROCESSING LETTERS, VOL. 13.
- [7] Wes Hatch, 2003, A Quick Review of Audio Fingerprinting, Retrieved Jan 17, 2011, from <http://www.music.mcgill.ca/~wes/docs/finger2.pdf>
- [8] Minho Jin and Chang D. Yoo. 2007. Temporal Dynamics for Spectral Sub-Band Centroid Audio Fingerprints, IEEE Multimedia and Expo.
- [9] Audible Magic. 2012, Technology Overview, Retrieved April 11, 2012, from: <http://www.audiblemagic.com/technology.php>
- [10] Bryan Jacobs. 2009. How Shazam Works. Retrieved Feb 1, 2011, from <http://laplacian.wordpress.com/2009/01/10/how-shazam-works/>
- [11] MusicTrace GmbH. 2005, Audio Fingerprinting, Retrieved April 11, 2012, from: <http://www.musictrace.de/technologies/fingerprinting.en.htm>
- [12] MusicTrace GmbH. 2005, Audio Watermarking, Retrieved April 11, 2012, <http://www.musictrace.de/technologies/watermarking.en.htm>