

เหมืองกระบวนการ Process Mining

วิเชียร เปรมชัยสวัสดิ์

บัณฑิตวิทยาลัย สาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยสยาม

E-mail: wichian@siam.edu

บทคัดย่อ

บทความนี้นำเสนอเทคนิคเหมืองกระบวนการซึ่งเป็นเทคโนโลยีอุบัติใหม่ที่มีวัตถุประสงค์เพื่อใช้ในการค้นหาคุณค่าจากข้อมูลที่เกิดขึ้นจริงในบันทึกเหตุการณ์ ปัจจุบันสารสนเทศและเทคโนโลยีการสื่อสารได้กระจายเข้าไปสู่ทุกภาคส่วนขององค์กร ระบบสารสนเทศเกิดขึ้นมาเพื่อตอบสนองต่อบริบทขององค์กรและกระบวนการธุรกิจ ทำให้มีข้อมูลพร้อมใช้จำนวนมากสำหรับการวิเคราะห์ต่างๆ บทความนี้กล่าวถึงภาพรวมของเทคนิคเหมืองกระบวนการ เครื่องมือที่ใช้ในการทำเหมืองกระบวนการ ข้อมูลที่ใช้ในการทำเหมืองกระบวนการ อัลกอริทึมสำคัญที่ใช้คือ α -algorithm พร้อมทั้งกรณีศึกษาในการประยุกต์ใช้เทคนิคเหมืองกระบวนการกับบันทึกเหตุการณ์ของระบบ e-learning ผลลัพธ์ของสารสนเทศที่ได้จากการทำเหมืองกระบวนการแสดงให้เห็นว่าสามารถนำมาใช้เพื่อทำให้เกิดความเข้าใจอย่างลึกซึ้งของกระบวนการธุรกิจที่ทำจริงๆ ในปัจจุบัน และสามารถนำมาใช้ในการออกแบบกระบวนการธุรกิจใหม่ เพื่อให้องค์กรมีกระบวนการธุรกิจที่มีประสิทธิภาพมากยิ่งขึ้น

คำสำคัญ : เหมืองกระบวนการ, α -algorithm

Abstract

The paper presents the process mining technique which is emerging technology aimed to discovery value of real data recorded in event log. Today, information and communication technology are distributed all over an organization. This information occurred to respond to the context of the organization and their business processes. Therefore, there is a huge of data ready available for various analysis. The paper describes the overall techniques of process mining, tools, data using in process mining. The α -algorithm which is an important algorithm is also presented. Finally, the case study of applying the process mining techniques for the event log of the e-learning system is presented. The resulted information got from process mining shows that it can be used to understand more insight of the present real processes. The information also can be used to redesign the new business process which has more efficiency for the organization.

1. บทนำ

ปัจจุบันการมีแข่งขันกันอย่างมากมายทั้งภายในประเทศและระดับนานาชาติ ก่อปรกกับการผลักดันนโยบายด้านเศรษฐกิจดิจิทัลของรัฐบาลในปัจจุบัน ทำให้เทคโนโลยีสารสนเทศมีผลกระทบต่อสังคมและธุรกิจ[1] ดังนั้นเพื่อความอยู่รอดในการแข่งขัน การทำงานให้มีประสิทธิภาพมากยิ่งขึ้นจึงเป็นสิ่งจำเป็นที่ไม่อาจหลีกเลี่ยงได้ การทำงานเป็นเรื่องของกระบวนการธุรกิจ ดังนั้นการทำงานให้มีประสิทธิภาพจึงต้องมีกระบวนการที่ดีและมีประสิทธิภาพ ในอดีตการออกแบบกระบวนการธุรกิจในองค์กรนั้นอาศัยเครื่องมือในการสร้างแบบจำลองและทดสอบโดยใช้ข้อมูลที่สร้างขึ้นจากความน่าจะเป็น ทำให้ได้แบบจำลองที่คาดว่าจะดำเนินการได้ แต่อย่างไรก็ตามในความเป็นจริงกระบวนการอาจไม่ได้ดำเนินการไปดังที่คาดหวัง ด้วยเครื่องมือที่มีอยู่อาจรับรู้ผลของกระบวนการทำงานที่ออกแบบไว้โดยสังเกตการณ์ การสอบถาม ซึ่งอาจได้รับทราบข้อเท็จจริงที่ยังไม่ตรงกับความเป็นจริงมากนัก

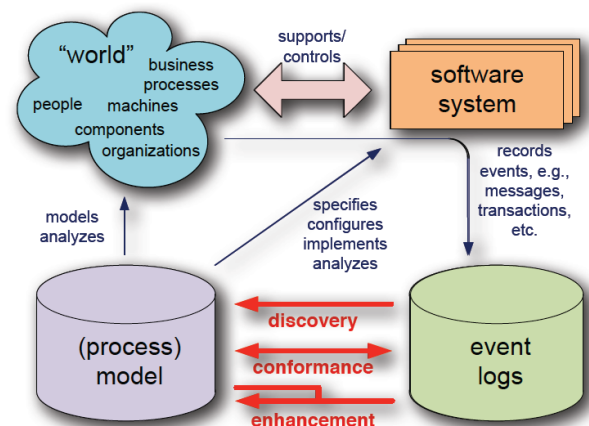
เหมืองกระบวนการ (Process mining) เป็นเทคนิคการวิเคราะห์ข้อมูลที่เน้นไปที่พฤติกรรมที่อยู่ในข้อมูลของบันทึกเหตุการณ์ (Event log) ดังนั้นเหมืองกระบวนการจึงเป็นเทคนิคที่สามารถนำเข้ามาเป็นเครื่องมือเพื่อช่วยในการแก้ปัญหาที่ทุกองค์กรส่วนใหญ่พบอยู่ก็คือ การที่มีสารสนเทศอย่างจำกัดในการที่จะบ่งบอกได้ว่ามีเหตุการณ์อะไรเกิดขึ้นจริงๆ ในองค์กร

บทความนี้จึงได้นำเสนอวิธีการในการค้นพบแบบจำลองกระบวนการธุรกิจ ด้วยเทคนิคเหมืองกระบวนการโดยใช้ข้อมูลที่เกิดขึ้นจริงในบันทึก

เหตุการณ์ไม่ใช่ข้อมูลที่สมมติขึ้น การนำเสนอของบทความนี้มีเค้าโครงดังนี้: ส่วนที่ 2 กล่าวถึงภาพรวมของเหมืองกระบวนการ ส่วนที่ 3 เป็นการแนะนำเครื่องมือที่ใช้ในการทำเหมืองกระบวนการ ส่วนที่ 4 กล่าวถึงข้อมูลที่ใช้ในการทำเหมืองกระบวนการ ได้แก่ บันทึกเหตุการณ์ ส่วนที่ 5 แนะนำอัลกอริทึมสำคัญที่ใช้ คือ α -algorithm ส่วนที่ 6 เป็นกรณีศึกษาในการประยุกต์ใช้เทคนิคเหมืองกระบวนการกับบันทึกเหตุการณ์ของระบบ e-learning และส่วนสุดท้ายจะเป็นการสรุปเนื้อหาในบทความนี้

2. เหมืองกระบวนการ

เหมืองกระบวนการ [2] เป็นเทคโนโลยีที่อุบัติขึ้นมาใหม่เมื่อไม่นานมานี้ เป็นเทคนิคที่ใช้ในการค้นหาคุณค่า (value) จากข้อมูลที่เกิดขึ้นจริงในบันทึกเหตุการณ์ โดยมีวัตถุประสงค์หลัก 3 ประการ คือ การค้นพบกระบวนการ (process discovery), การตรวจสอบความสอดคล้อง (conformance checking), และ การปรับปรุงให้ดีขึ้น (enhancement) ภาพรวมการทำงานของการทำงานเหมืองกระบวนการได้แสดงในรูปที่ 1

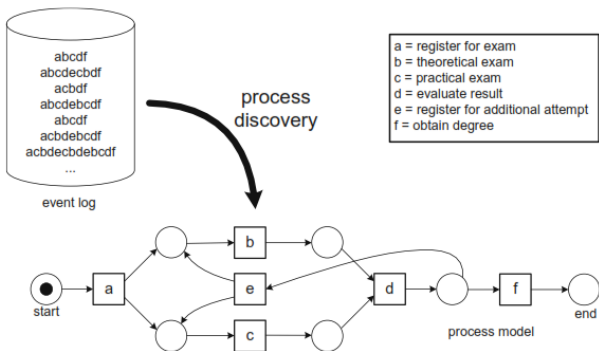


รูปที่ 1 ภาพรวมของการทำเหมืองกระบวนการ [1]

ผลลัพธ์ของการทำเหมืองกระบวนการทำให้ผู้ใช้งานสามารถมองเห็นความสัมพันธ์ระหว่างกิจกรรมต่างๆ ของข้อมูลเหตุการณ์อย่างลึกซึ้งตั้งแต่ต้นกระบวนการจนจบกระบวนการธุรกิจ จึงทำให้ได้รับความสนใจเพิ่มมากขึ้นจากทั้งในภาคอุตสาหกรรม นักวิจัย นักพัฒนา และยังได้รับการสนับสนุนจาก IEEE Task Force on Process mining [3] ในปี ค.ศ. 2009 อีกด้วย

2.1 การค้นพบกระบวนการ (process discovery)

เป็นเทคนิคที่ได้รับความนิยมมากที่สุด ใช้ในการสร้างแบบจำลองกระบวนการ (process model) โดยใช้บันทึกเหตุการณ์เป็นข้อมูลเข้าและไม่ต้องมีสารสนเทศอื่นๆล่วงหน้า ผลลัพธ์ที่ได้คือแบบจำลองกระบวนการธุรกิจ โดยทั่วไปจะนำเสนอในรูปแบบของ Petri Net [4] ดังแสดงในรูปที่ 2



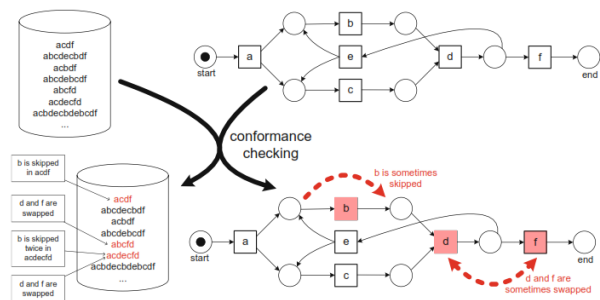
รูปที่ 2 การค้นพบกระบวนการ [5]

เป้าหมายของการค้นพบแบบจำลองกระบวนการอยู่ที่การสร้างแบบจำลองในระดับของ Petri Net มากกว่าที่จะเป็นตัวแทนในอย่างลำดับ (sequential) หรือตัวแทนในระดับที่ต่ำกว่านี้ เช่น

Markov chains, Finite state machine หรือ Regular expression ดังนั้นเทคนิคของเหมืองกระบวนการจึงต้องมีความสามารถในการจัดการกับการทำงานพร้อมๆกัน (concurrency) ในรูปแบบต่างๆกันได้ด้วย

2.2 การตรวจสอบความสอดคล้อง (conformance checking)

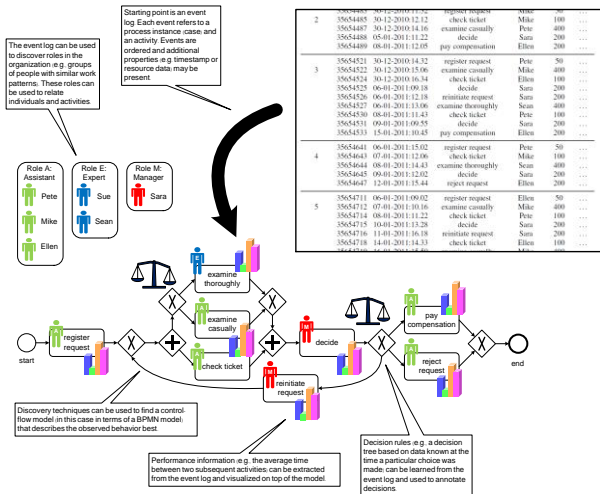
เป็นเทคนิคที่ใช้ในการเปรียบเทียบแบบจำลองกระบวนการที่มีอยู่กับพฤติกรรมจริงที่ได้บันทึกไว้ในบันทึกเหตุการณ์ว่ามีความสอดคล้องกันมากน้อยเพียงใด มีส่วนใดบ้างที่เบี่ยงเบนไปจากแบบจำลองกระบวนการธุรกิจที่กำหนดไว้ ดังแสดงในรูปที่ 3



รูปที่ 3 การตรวจสอบความสอดคล้อง [5]

2.3 การปรับปรุงให้ดีขึ้น (enhancement)

เป็นการขยายขีดความสามารถของแบบจำลองกระบวนการที่มีอยู่ที่เน้นไปในเรื่องของการไหลของการควบคุมการ (control flow) ไปในมุมมองอื่นๆ เช่น มุมมองด้านองค์กร เวลา เครือข่ายสังคม เป็นต้น หรือ การออกแบบแบบจำลองใหม่เพื่อแก้ปัญหาต่างๆในแบบจำลองเดิม เช่น ปัญหาคอขวด เป็นต้น ดังแสดงในรูปที่ 4



รูปที่ 4 การปรับปรุงให้ดีขึ้น [2]

3. เครื่องมือ

ปัจจุบันมีเครื่องมือจำนวนมากที่สนับสนุนการทำเหมืองกระบวนการทั้งในรูปแบบของ open source และเชิงพาณิชย์ ในบทความนี้จะแนะนำตัวอย่างที่มาจากทั้งสองค่ายที่ได้รับความนิยมอย่างมากและสามารถนำมาทดลองใช้ได้

3.1 ProM [6]

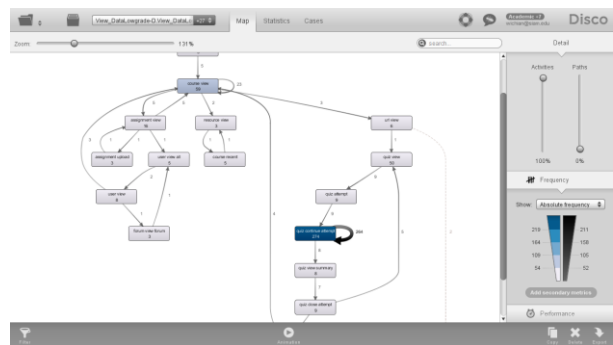
ProM (<http://www.processmining.org>) เป็นตัวอย่างของซอฟต์แวร์ที่เป็น open source ที่มีชุดของเครื่องมือที่ใช้ในการสนับสนุนเทคนิคในการทำเหมืองกระบวนการที่หลากหลายโดยการสนับสนุนสถาปัตยกรรมแบบ Plug-in ทำให้นักวิจัยสามารถเพิ่มอัลกอริทึมใหม่ๆเข้าไปเพิ่มเติมได้ ProM ใช้ข้อมูลเข้าในรูปแบบของ MXML, SA-MXML, หรือ XES รุ่นล่าสุดของ ProM อยู่ที่ version 6.4.1 ดังแสดงในรูปที่ 5



รูปที่ 5 ProM 6.4.1

3.2 Disco [7]

เป็นเครื่องมือเชิงพาณิชย์ที่สนับสนุนเทคนิคในการทำเหมืองกระบวนการที่เป็นมิตรกับผู้ใช้ มีความสะดวกในการใช้และเรียนรู้โดยไม่จำเป็นต้องมีประสบการณ์ในการทำเหมืองกระบวนการ แต่สนับสนุนการทำงานหลักๆเฉพาะ Fuzzy miner (<https://fluxicon.com/disco/>) ดังแสดงในรูปที่ 6 ข้อมูลเข้าในรูปแบบของ MXML, SA-MXML, หรือ XES และที่สำคัญอีกประการหนึ่งคือ Disco สามารถรับข้อมูลที่อยู่ในรูปแบบของ csv ได้ ทำให้มีความสะดวกในการใช้งาน พร้อมทั้งสามารถ export ข้อมูลออกไปเพื่อใช้ในการวิเคราะห์ด้วยอัลกอริทึมอื่นๆที่ ProM สามารถทำการวิเคราะห์ให้ได้อีกด้วย



รูปที่ 6 Disco

4. บันทึกเหตุการณ์

บันทึกเหตุการณ์ (Event Log) คือ ชุดของเหตุการณ์ใช้เพื่อเป็นข้อมูลเข้าของการทำเหมืองกระบวนการ ดังตัวอย่างที่แสดงในรูปที่ 7

Time	IPAddress	Case	Action	Grade
12/1/2013	183.89.250.225	Nathap	user view	A
25/12/2012	171.100.10.197	Jirayu	M course view	A
25/12/2012	171.100.10.197	Jirayu	M quiz view	A
17/12/2012	115.87.92.57	SALEE	F course view	B
14/12/2012	115.31.135.226	SALEE	F course view	B
14/12/2012	110.164.218.150	Nichapa	course view	A
14/12/2012	110.164.218.150	Nichapa	course view	A
14/12/2012	124.122.109.38	Tanasal	course view	B+
14/12/2012	124.122.109.38	Tanasal	course view	B+
14/12/2012	124.122.109.38	Tanasal	course view	B+
14/12/2012	124.122.109.38	Tanasal	url view	B+
14/12/2012	124.122.109.38	Tanasal	url view	B+
14/12/2012	124.122.109.38	Tanasal	course view	B+
14/12/2012	124.122.109.38	Tanasal	course view	B
14/12/2012	124.122.92.99	SALEE	F course view	B
14/12/2012	124.122.92.99	SALEE	F course view	B
14/12/2012	125.24.244.29	Sompol	course view	A
14/12/2012	124.122.92.99	SALEE	F url view	B
14/12/2012	124.120.39.158	Nucha	url view	B+
14/12/2012	124.120.39.158	Nucha	url view	B+
14/12/2012	125.24.244.29	Sompol	course view	A
14/12/2012	124.120.39.158	Nucha	url view	B+
14/12/2012	124.120.39.158	Nucha	url view	B+
14/12/2012	125.24.244.29	Sompol	url view	A
14/12/2012	124.122.92.99	SALEE	F course view	B
14/12/2012	124.122.92.99	SALEE	F quiz view	B
14/12/2012	110.171.63.166	Jirayu	M course view	A

รูปที่ 7 ตัวอย่างของข้อมูลในบันทึกเหตุการณ์

โดยที่เหตุการณ์ไม่มีความจำเป็นที่ถูกจัดเก็บไว้ในแฟ้มบันทึกที่แยกออกมาโดยเฉพาะ เช่น เหตุการณ์อาจจะถูกบันทึกกระจายอยู่ยู่ในตารางที่ต่างกันในฐานะข้อมูล โดยทั่วไปบันทึกเหตุการณ์จะประกอบไปด้วยสารสนเทศเกี่ยวกับเหตุการณ์ที่อ้างถึงกิจกรรม (activity) ต่างๆที่เกิดขึ้น และ กรณี (case) โดยที่ กรณี หรือ บางครั้งเรียกว่ากระบวนการ ณ ขณะนั้น (process instance) คือ สิ่งที่เราต้องการบริหารจัดการนั่นเอง เช่น คำสั่งซื้อของลูกค้า ผู้ป่วยในโรงพยาบาล เป็นต้น ส่วนกิจกรรมหมายถึงการดำเนินการที่มีในกรณีนั้นๆ นอกจากนั้นแล้วบันทึกเหตุการณ์มักจะประกอบด้วยประทับเวลา

(timestamp) เพื่อเป็นการบ่งชี้เวลาที่เหตุการณ์ต่างๆ เกิดขึ้น และเมื่อมีบุคคลเข้ามาเกี่ยวข้องกับ บันทึกเหตุการณ์มักจะมีการบันทึกบุคคลที่กระทำในแต่ละกิจกรรมไว้ด้วย ซึ่งข้อมูลเหล่านี้สามารถใช้ในการทำเหมืองกระบวนการเพื่อหาเครือข่ายสังคม (social network)

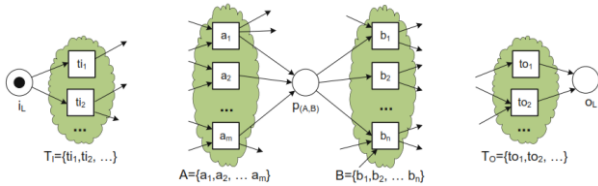
5. อัลกอริทึมที่ใช้ในการค้นพบกระบวนการ

มีอัลกอริทึมจำนวนมากที่สามารถนำมาใช้ในการค้นพบกระบวนการ เช่น α -algorithm, region-based approach และการใช้ genetic algorithms [8][9] ในบทความนี้จะแนะนำเฉพาะ α -algorithm [2] เนื่องจากเป็นอัลกอริทึมพื้นฐานที่จะสามารถทำให้เกิดความเข้าใจและสามารถนำไปประยุกต์ใช้ แต่เป็นอัลกอริทึมที่มีข้อจำกัดในการจัดการกับสัญญาณรบกวน (noise) การทำซ้ำ (loop) เป็นต้น α -algorithm มีขั้นตอนการทำงานดังนี้

- $T_L = \{t \in A \mid \exists \sigma \in L, t \in \sigma\}$,
- $T_I = \{t \in A \mid \exists \sigma \in L, t = \text{first}(\sigma)\}$,
- $T_O = \{t \in A \mid \exists \sigma \in L, t = \text{last}(\sigma)\}$,
- $X_L = \{(A, B) \mid A \subseteq T_L \wedge A \neq \emptyset \wedge B \subseteq T_L \wedge B \neq \emptyset \wedge \forall a \in A \forall b \in B, a \rightarrow_L b \wedge \forall a_1, a_2 \in A, a_1 \#_L a_2 \wedge \forall b_1, b_2 \in B, b_1 \#_L b_2\}$,
- $Y_L = \{(A, B) \in X_L \mid \forall (A', B') \in X_L, A \subseteq A' \wedge B \subseteq B' \Rightarrow (A, B) = (A', B')\}$,
- $P_L = \{p_{(A, B)} \mid (A, B) \in Y_L\} \cup \{(i_L, o_L)\}$,
- $F_L = \{(a, p_{(A, B)}) \mid (A, B) \in Y_L \wedge a \in A\} \cup \{(p_{(A, B)}, b) \mid (A, B) \in Y_L \wedge b \in B\} \cup \{(i_L, t) \mid t \in T_I\} \cup \{(t, o_L) \mid t \in T_O\}$,
- $l_L \in T_L \rightarrow A$ with $l(t) = t$ for $t \in T_L$, and
- $\alpha(L) = (N, M_{init}, M_{final})$ with $N = (P_L, T_L, F_L, l_L)$, $M_{init} = [i_L]$, $M_{final} = [o_L]$.

การทำงานของอัลกอริทึมอธิบายได้ดังนี้ ขั้นตอนที่ 1 เป็นการตรวจสอบว่ากิจกรรม (activity) อยู่ในบันทึก (log) T_L ขั้นตอนที่ 2 กำหนด set ของ start activities ขั้นตอนที่ 3 กำหนด set ของ end

activities ขั้นตอนที่ 4 และ 5 เป็นการกำหนด place ใน Petri Net และการเชื่อมต่อ ในการสร้าง place, $p_{(A,B)}$ โดยที่ A เป็น set ของ input transition ($\bullet p_{(A,B)} = A$) และ B เป็น set ของ output transition ($p_{(A,B)} \bullet = B$) ของ $p_{(A,B)}$



รูปที่ 8 การหา $p_{(A,B)}$ และการเชื่อมต่อ [5]

ขั้นตอนที่ 6 เป็นการเพิ่ม source place i_L และ sink place o_L ขั้นตอนที่ 7 ลากเส้นเชื่อมโยงใน Petri Net ขั้นตอนที่ 8 และ 9 ได้ Petri Net ที่สมบูรณ์

6. กรณีศึกษา

จากหัวข้อที่ได้กล่าวมาแล้วได้ให้ความรู้พื้นฐานในการทำเหมืองกระบวนการ ในส่วนนี้จะได้นำเทคนิคเหมืองกระบวนการไปใช้ในการวิจัย โดยนำไปใช้ในการวิเคราะห์บันทึกเหตุการณ์ที่ได้จากระบบ e-learning ที่ใช้ Moodle [10] ซึ่งเป็น open source learning management system ถึงแม้ว่า Moodle จะมีความสามารถที่จะสร้างรายงานได้หลายรูปแบบ แต่ก็ไม่สามารถแสดงรายละเอียดในระดับกระบวนการที่จะทำให้อาจารย์ผู้สอนในรายวิชานั้นๆ เข้าใจถึงพฤติกรรมของนักศึกษาที่ใช้ระบบ โดยเฉพาะอย่างยิ่งก็คือกิจกรรมใดที่มีผลกระทบต่อผลสัมฤทธิ์ของนักศึกษา การวิจัยนี้จึงใช้เทคนิคเหมืองกระบวนการในการหาคำตอบนี้

6.1 บันทึกเหตุการณ์

เนื่องวัตถุประสงค์ของงานวิจัยนี้ต้องการค้นหาว่ากิจกรรมใดที่มีผลกระทบต่อผลสัมฤทธิ์ของนักศึกษา ดังนั้นจึงได้แบ่งข้อมูลของนักศึกษาออกเป็น 2 กลุ่มคือ กลุ่มที่มีผลสัมฤทธิ์สูง คือได้เกรด A, B+ และ B กับกลุ่มที่มีผลสัมฤทธิ์ต่ำ คือได้เกรด C+ และต่ำกว่า เพื่อสร้างแบบจำลองกระบวนการของทั้งสองกลุ่มและทำการเปรียบเทียบรายละเอียดของพฤติกรรมของนักศึกษาในการทำกิจกรรมในแบบจำลองกระบวนการทั้งสอง ตัวอย่างของบันทึกเหตุการณ์ที่ใช้ในการวิจัยนี้อยู่ในรูปแบบของ csv ดังแสดงไว้ในรูปที่ 9 ที่ใช้ในการวิเคราะห์โดยโปรแกรม Disco และรูปที่ 10 โดยโปรแกรม ProM 6.4.1 ในงานวิจัยนี้จะใช้โปรแกรม Disco ในการวิเคราะห์เท่านั้น โดยรูปที่ 11 แสดงภาพรวมแบบจำลองกระบวนการที่ได้จากบันทึกเหตุการณ์ในรูปที่ 9 ของนักศึกษาทั้งสองกลุ่มจำนวน 37 คน

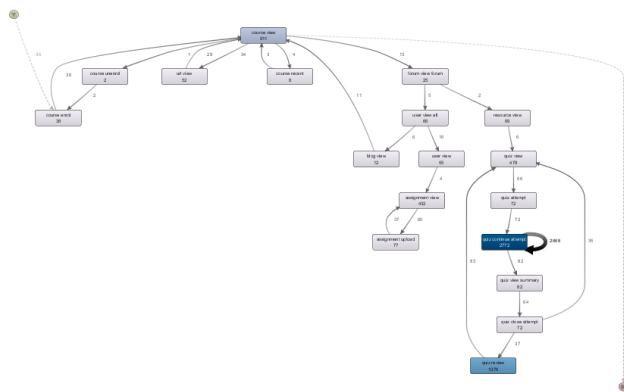
Time	IPaddress	name	Action	Grade
25580112 11:35:32	183.89.250.225	Nathap...	user view	A
25551225 14:34:59	171.100.10.197	Jirayu M...	course view	A
25551225 14:33:21	171.100.10.197	Jirayu M...	quiz view	A
25551217 16:57:36	115.87.92.57	SALEE ...	course view	B
25551214 18:19:37	115.31.135.226	SALEE ...	course view	B
25551214 17:01:00	110.164.218.150	Nichap...	course view	A
25551214 17:00:49	110.164.218.150	Nichap...	course view	A
25551214 16:55:53	124.122.109.38	Tanasa...	course view	B+
25551214 16:55:49	124.122.109.38	Tanasa...	course view	B+
25551214 16:55:44	124.122.109.38	Tanasa...	course view	B+
25551214 16:55:30	124.122.109.38	Tanasa...	course view	B+
25551214 16:50:05	124.122.109.38	Tanasa...	url view	B+
25551214 16:45:03	124.122.109.38	Tanasa...	url view	B+
25551214 16:44:48	124.122.109.38	Tanasa...	course view	B
25551214 16:38:59	124.122.92.99	SALEE ...	course view	B
25551214 16:37:16	124.122.92.99	SALEE ...	course view	B
25551214 16:34:25	125.24.244.29	Sompol...	course view	A
25551214 16:32:30	124.122.92.99	SALEE ...	url view	B
25551214 16:16:16	124.120.39.158	Nucha...	url view	B+
25551214 16:16:04	124.120.39.158	Nucha...	course view	B+
25551214 16:14:39	125.24.244.29	Sompol...	course view	A
25551214 16:14:16	124.120.39.158	Nucha...	url view	B+
25551214 16:13:45	124.120.39.158	Nucha...	course view	B+
25551214 16:09:32	125.24.244.29	Sompol...	url view	A
25551214 16:08:35	124.122.92.99	SALEE ...	course view	B
25551214 16:08:25	124.122.92.99	SALEE ...	quiz view	B
25551214 16:01:01	110.171.83.166	Jirayu M...	course view	A
25551214 15:42:03	110.168.75.89	Drawan...	url view	B
25551214 15:41:48	110.168.75.89	Drawan...	url view	B

รูปที่ 9 ตัวอย่างของบันทึกเหตุการณ์ที่นำเข้าด้วยโปรแกรม

Disco



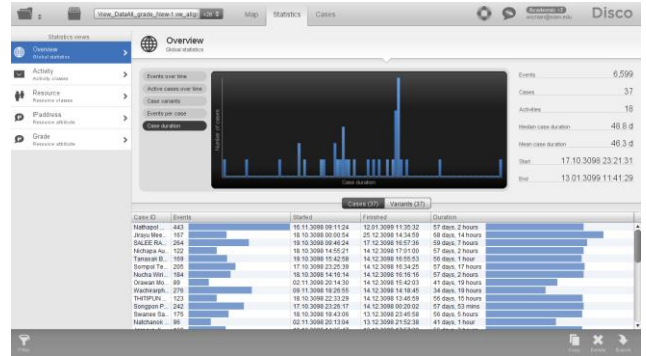
รูปที่ 10 ตัวอย่างของบันทึกเหตุการณ์ที่นำเข้ามาด้วยโปรแกรม ProM 6.4.1



รูปที่ 11 แบบจำลองกระบวนการที่ได้จากบันทึกเหตุการณ์ในรูปที่ 9

6.2 ผลลัพธ์ของเหมืองกระบวนการและการค้นพบ

จากค่าทางสถิติที่แสดงในรูปที่ 12 และรูปที่ 13 แสดงให้เห็นว่ามีเหตุการณ์ที่เกิดขึ้นทั้งหมด 6,599 เหตุการณ์, 37 กรณี (จำนวนนักศึกษา) และ 18 กิจกรรม ระยะเวลาที่ใช้โดยเฉลี่ยเท่ากับ 46.3 วัน



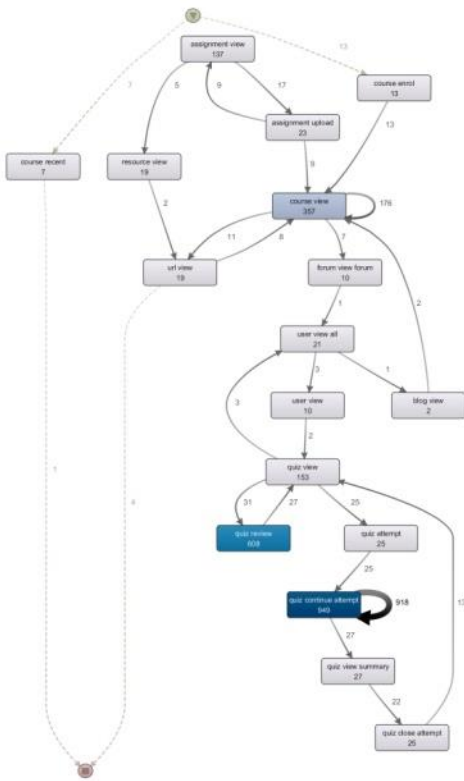
รูปที่ 12 ข้อมูลทางสถิติของบันทึกเหตุการณ์ในรูปที่ 9

Events	6,599
Cases	37
Activities	18
Median case duration	48.8 d
Mean case duration	46.3 d
Start	17.10.3098 23:21:31
End	13.01.3099 11:41:29

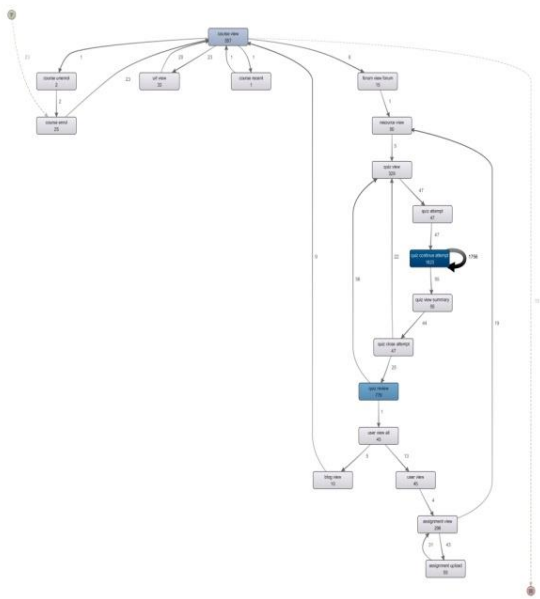
รูปที่ 13 ข้อมูลทางสถิติจากรูปที่ 12

ในการทำเหมืองกระบวนการของกลุ่มที่มีผลสัมฤทธิ์สูงที่มีนักศึกษาจำนวน 24 คน และกลุ่มที่มีผลสัมฤทธิ์ต่ำที่มีนักศึกษาจำนวน 13 คน พบว่ากลุ่มที่มีผลสัมฤทธิ์สูงใช้เวลาเรียนเฉลี่ย 47.4 วัน ซึ่งมากกว่ากลุ่มที่มีผลสัมฤทธิ์ต่ำที่ใช้เวลาเรียนเฉลี่ย 44.2 วัน นอกจากนี้ กลุ่มที่มีผลสัมฤทธิ์สูงมีการเข้ามาทำกิจกรรมต่างๆในระบบ 4,216 ครั้ง ในขณะที่กลุ่มที่มีผลสัมฤทธิ์ต่ำมีกิจกรรม 2,383 ครั้ง แต่เมื่อพิจารณาจากค่าเฉลี่ยแล้วกลับพบว่า กลุ่มที่มีผลสัมฤทธิ์ต่ำมีกิจกรรมกับระบบสูงกว่ากลุ่มที่มีผลสัมฤทธิ์สูงคือ 183 (2384/13) ต่อ 176 (4216/24) ครั้ง แบบจำลองกระบวนการของทั้งสองกลุ่มที่กล่าวมานี้โดยใช้ Fuzzy model ซึ่งวิเคราะห์จากความถี่

ของกิจกรรมที่กระทำโดยใช้โปรแกรม Disco ได้แสดงไว้ในรูปที่ 14 และ 15

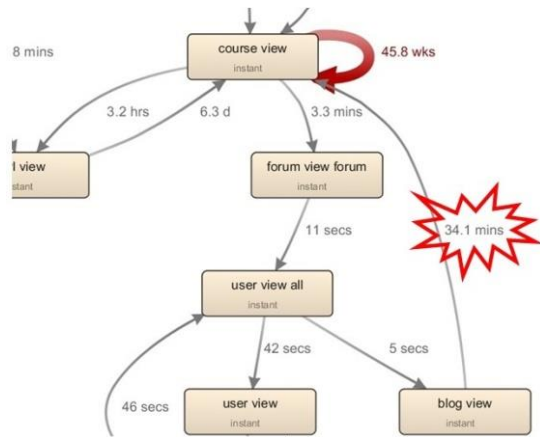


รูปที่ 14 Fuzzy model โดยใช้ความถี่ของกิจกรรมของกลุ่มที่มีผลสัมฤทธิ์ต่ำ

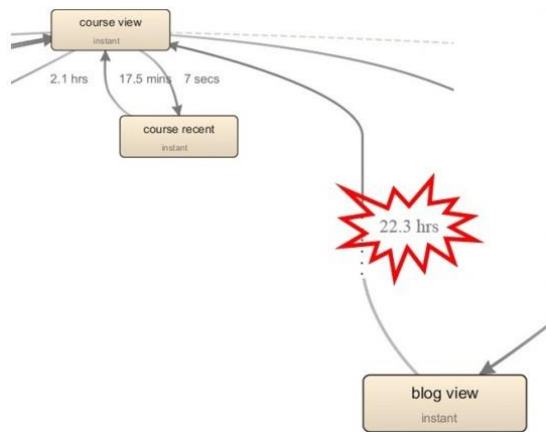


รูปที่ 15 Fuzzy model โดยใช้ความถี่ของกิจกรรมของกลุ่มที่มีผลสัมฤทธิ์สูง

จะเห็นได้ว่าการพิจารณาจากความถี่ของกิจกรรมไม่ได้ให้สารสนเทศที่สะท้อนถึงผลสัมฤทธิ์ของนักศึกษามากนัก งานวิจัยนี้จึงได้ Fuzzy model โดยใช้ระยะเวลาที่ใช้ของกิจกรรมเพื่อทำการเปรียบเทียบดังแสดงในรูปที่ 16 และ 17



รูปที่ 16 Fuzzy model โดยใช้ระยะเวลาที่ใช้ของกิจกรรมของกลุ่มที่มีผลสัมฤทธิ์ต่ำ



รูปที่ 17 Fuzzy model โดยใช้ระยะเวลาที่ใช้ของกิจกรรมของกลุ่มที่มีผลสัมฤทธิ์สูง

จากการใช้ Fuzzy model โดยใช้ระยะเวลาที่ใช้ของกิจกรรม พบว่ากลุ่มที่มีผลสัมฤทธิ์ต่ำใช้เวลา

เพียง 34.1 นาทีในขณะที่กลุ่มที่มีผลสัมฤทธิ์สูงใช้เวลาถึง 22.3 ชั่วโมง

จากการศึกษาจากกรณีศึกษาที่ได้กล่าวถึงข้างต้น โดยใช้โปรแกรม Disco โดยใช้ระยะเวลาที่ใช้และความถี่ของกิจกรรม ของ Fuzzy model พบว่า นักศึกษามีความถี่ในการทำกิจกรรมไม่แตกต่างกันมากนัก แต่ระยะเวลาที่ใช้ในการดูและอ่านใน “blog page” มีความแตกต่างกันมากในกลุ่มที่มีผลสัมฤทธิ์สูงกับกลุ่มที่มีผลสัมฤทธิ์ต่ำ การเข้าใจถึงกิจกรรมต่างๆที่นักศึกษาใช้ทำให้ผู้สอนสามารถนำมาใช้เพื่อปรับกระบวนการในการสอนให้ดีขึ้นได้ ซึ่งข้อมูลเหล่านี้ยากที่จะได้มาจากวิธีการอื่น พร้อมทั้งการอธิบายด้วยภาพทำให้ผู้ใช้เข้าใจได้ง่ายขึ้น

7. สรุปและงานในอนาคต

บทความฉบับนี้ได้นำเสนอเทคนิคของเหมืองกระบวนการเพื่อใช้ในการค้นพบแบบจำลองกระบวนการธุรกิจที่ปฏิบัติกันจริงๆในองค์กร โดยใช้ข้อมูลจากบันทึกเหตุการณ์ที่เกิดขึ้นจริงและส่วนใหญ่มีอยู่แล้วในระบบสารสนเทศขององค์กร แบบจำลองกระบวนการธุรกิจที่ค้นพบ ทำให้องค์กรเข้าใจกระบวนการทำงานที่แท้จริงขององค์กรอย่างลึกซึ้งสามารถนำสารสนเทศที่ได้มาใช้ในการปรับปรุงกระบวนการธุรกิจให้ดีและมีประสิทธิภาพมากยิ่งขึ้นอันจะเป็นการเพิ่มขีดความสามารถในการแข่งขันในโลกปัจจุบัน นอกจากนี้บทความนี้ยังได้นำเสนอกรณีศึกษาของการประยุกต์ใช้เหมืองกระบวนการกับการศึกษา เพื่อค้นหากระบวนการในการศึกษาของนักศึกษาที่มีต่อระบบออนไลน์ เทคนิคเหมืองกระบวนการให้แนวทางใหม่ในการวิเคราะห์ข้อมูลใน

ระดับกระบวนการตั้งแต่ต้นกระบวนการจนจบกระบวนการ ทำให้เกิดความเข้าใจอย่างลึกซึ้งในพฤติกรรมในกิจกรรมที่เกิดขึ้นจริง ที่ไม่อาจค้นพบด้วยเทคนิคอื่น

นอกจากการประยุกต์ใช้เทคนิคในการทำเหมืองกระบวนการดังที่กล่าวไปแล้ว ซึ่งเป็นการวิเคราะห์กระบวนการโดยเน้นที่การไหลของการควบคุม ในอนาคตจะได้ขยายการวิจัยออกไปโดยใช้อัลกอริทึมแบบอื่นๆ ตลอดจนการใช้มุมมองอื่นๆ เช่น มุมมองด้านเครือข่ายสังคมเป็นต้น นอกจากนี้ยังมีความท้าทายในการวิจัยเพื่อสร้างอัลกอริทึมใหม่ หรือการปรับปรุงกระบวนการให้มีประสิทธิภาพมากยิ่งขึ้น โดยเฉพาะอย่างยิ่งการทำเหมืองกระบวนการกับข้อมูลบันทึกเหตุการณ์แบบออนไลน์ เพื่อให้สามารถนำมาสนับสนุนการตัดสินใจแบบเวลาจริง (real time)

เอกสารอ้างอิง

- [1] N. Premchaiswadi. The Impact of Information Technology on Society and Business, Engineering Journal of Siam University, Vol.13, Issue 1, No.24, 2012, pp.50-67.
- [2] W. van der Aalst. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer-Verlag, Berlin, 2011.
- [3] IEEE Task Force on Process Mining, Process Mining Manifesto. In BPM Workshops, volume99 of Lecture Notes in

- Business Information Processing,
SpringerVerlag, Berlin, 2011, 169–194.
- [4] K. Jensen, L. Kristensen, Coloured Petri
Nets: Modelling and Validation of Concurrent
Systems, Springer, 2009.
- [5] W. van der Aalst. Process Mining in the
Large: A Tutorial, E. Zimanyi (Ed.): eBISS
2013, LNBIP 172, pp. 33–76, 2014. DOI:
10.1007/978-3-319-05461-2_2, Springer
International Publishing Switzerland 2014
- [6] ProM. <http://www.processmining.org>
- [7] Fluxicon. <https://fluxicon.com/disco/>
- [8] Buijs, J.C.A.M., van Dongen, B.F., van der
Aalst, W.M.P.: On the role of fitness,
precision, generalization and simplicity in
process discovery. In: Meersman, R., et al.
(eds.) OTM 2012, Part I. LNCS, vol. 7565,
pp. 305–322. Springer, Heidelberg (2012).
- [9] Alves de Medeiros, A.K., Weijters, A.J.M.M.,
van der Aalst, W.M.P.: Genetic process
mining: an experimental evaluation. *Data
Min. Knowl. Disc.* 14(2), 245– 304 (2007).
- [10] van der Aalst, W.M.P., Weijters, A.J.M.M.,
Maruster, L.: Workflow mining: discovering
process models from event logs. *IEEE
Trans. Knowl. Data Eng.* 16(9), 1128–1142
(2004).
- [11] Moodle. <https://moodle.org/>