

PENERAPAN DATA MINING UNTUK ANALISIS POLA PEMBELIAN KONSUMEN DENGAN ALGORITMA FP-GROWTH PADA DATA TRANSAKSI PENJUALAN SPARE PART MOTOR

Alfannisa Annurullah Fajrin¹, Algifanri Maulana²

Teknik Informatika, Universitas Putera Batam
Jalan R. Soeprapto, Muka Kuning – Batam
asykharit1302@gmail.com, algifm@gmail.com

Abstract

The company has many branch or dealer as CV. This new TJAJAJA, requires locations that must be considered in setting up a new branch, should be able to influence consumer purchasing patterns, because the pattern of purchase every consumer is different – different. This needs to be analyzed further so that it can yield useful information, as well as maximize the benefits to be gained. Data Mining can be used by large companies to dig up the data to get information that can support and improve the business processes of the company. Then in this research done testing with FP-Growth algorithm to help companies figure out the pattern of consumer purchase transactions and sales of spare parts.

Keywords: Data Mining, Algoritma FP-Growth, Consumer Purchasing

Abstrak

Pada perusahaan yang mempunyai banyak cabang atau dealer seperti CV.TJAJAJA BARU ini, membutuhkan lokasi yang harus diperhatikan dalam mendirikan sebuah cabang baru, harus dapat mempengaruhi pola pembelian konsumen, karena pola pembelian setiap konsumen berbeda – beda. Hal ini perlu dianalisis lebih jauh sehingga dapat menghasilkan informasi yang bermanfaat, serta memaksimalkan keuntungan yang bisa diperoleh. Data Mining bisa digunakan oleh perusahaan besar untuk menggali data untuk mendapatkan informasi yang dapat menunjang dan meningkatkan proses bisnis perusahaan tersebut. Maka dalam penelitian ini dilakukan pengujian menggunakan algoritma FP-Growth untuk membantu perusahaan mengetahui pola pembelian konsumen dan transaksi penjualan spare part.

Kata kunci: Data Mining, Algoritma FP-Growth, Pembelian konsumen

1. PENDAHULUAN

Setiap perusahaan mempunyai data yang tersimpan dalam basis datanya. Data transaksi tersebut semakin hari semakin banyak dan bertambah. Dengan

bertambahnya jumlah data pada perusahaan tersebut, maka peran analisis untuk menganalisis data secara manual perlu digantikan dengan aplikasi yang berbasis komputer. Sehingga proses penganalisis dapat dilakukan secara tepat dan akurat.

Penelitian ini menerapkan algoritma FP-Growth dalam aplikasi yang dapat mengetahui pola pembelian konsumen pada tiap – tiap cabang yang berbeda dengan karakteristik yang berbeda pula. Dari pola yang dihasilkan tersebut akan didapatkan sebuah informasi.

Analisis asosiasi dikenal juga sebagai salah satu teknik data mining yang menjadi dasar dari berbagai teknik data mining lainnya. Khususnya salah satu tahap dari analisis asosiasi yang disebut analisis pola frekuensi tinggi (*frequent pattern mining*) menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien. Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, *support* (nilai penunjang) yaitu persentase kombinasi item tersebut dalam database dan *confidence*.

Dalam penelitian ini, algoritma yang digunakan adalah *Frequent Pattern-Growth (FP-Growth)* yaitu pengembangan dari metode *Apriori* yang merupakan salah satu alternatif untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data dengan membangkitkan struktur data *Tree* atau disebut dengan *Frequent Pattern Tree (FP-Tree)* [3].

Berdasarkan latar belakang yang diuraikan, maka terdapat beberapa hal yang menjadi permasalahan penelitian, yaitu : (1) Bagaimana menerapkan metode *Data Mining* dengan algoritma FP-Growth kedalam aplikasi untuk analisis pola pembelian konsumen ? (2) Bagaimana pola pembelian konsumen pada TB – Damar ? (3) Bagaimana interpretasi dari pola pembelian yang dihasilkan menjadi sebuah informasi ?

Dari rumusan masalah yang telah diuraikan di atas, ada beberapa tujuan dari penelitian ini, yaitu mengetahui pola pembelian konsumen pada masing – masing cabang, menginterpretasikan pola yang telah dihasilkan menjadi sebuah informasi, merancang bagaimana strategi dan sistem penjualan yang baik, menerapkan metode *Data Mining* dengan algoritma FP-Growth ke dalam aplikasi analisis pola pembelian konsumen, dan menguji pola yang telah diterapkan berhasil atau tidak.

2. METODOLOGI PENELITIAN

2.1. Metode Pengumpulan Data

Dalam penelitian ini menggunakan beberapa metode dalam pengumpulan data, yaitu :

a. Penelitian lapangan (Field Research)

Penelitian yang dilakukan langsung ke Mendi Shopping untuk memperoleh data primer. Adapun teknik pengumpulan data yang digunakan adalah :

- 1) Observasi, yaitu mengumpulkan data dengan melakukan pengamatan langsung dan pencatatan secara sistematis dengan tujuan agar memperoleh data yang objektif.
- 2) Wawancara langsung (interview), yaitu metode pengumpulan data dengan melakukan sesi tanya jawab secara langsung dengan pihak-pihak yang bersangkutan dalam bidang yang diteliti untuk memperoleh informasi yang dibutuhkan.

b. Penelitian perpustakaan (Library Research)

Peneliti melakukan penelitian dengan melakukan tinjauan ke pustaka guna mempelajari beberapa buku yang berhubungan dengan penelitian ini. Adapun beberapa buku tersebut diantaranya adalah buku-buku yang berhubungan dengan data mining dan algoritma FP Growth, buku - buku mengenai data transaksi penjualan dan juga buku - buku lainnya yang dirasa perlu dalam penunjangan penelitian ini.

c. Penelitian Laboratorium (Laboratory Research)

Merupakan tahap penelitian yang dilakukan dengan cara research labor komputer guna mempraktekkan langsung hasil dari analisa yang bertujuan untuk menguji kebenaran sistem yang dirancang.

Teori-teori yang menjadi landasan dalam penulisan tesis ini antara lain teori KDD, Data Mining, Association Rule, dan Algoritma FP Growth.

2.2. Knowledge Discovery in Database (KDD)

Knowledge discovery in Database (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses *knowledge discovery* melibatkan hasil dari proses *Data Mining* (proses mengekstrak kecenderungan pola suatu data), kemudian mengubah hasilnya se cara akurat menjadi informasi yang mudah dipahami. KDD sendiri diartikan sebagai keseluruhan proses non-trivial untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti.

2.3. Data Mining

Data Mining bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan *Data Mining* adalah kenyataan bahwa *Data Mining* mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu. Berawal dari beberapa disiplin ilmu, *Data Mining* bertujuan untuk memperbaiki teknik tradisional sehingga bisa menangani:

- a. Jumlah data yang sangat besar
- b. Dimensi data yang tinggi
- c. Data yang heterogen dan berbeda sifat

Pengelompokan *Data Mining* dibagi menjadi beberapa kelompok [1] yaitu :

- a. Deskripsi
Deskripsi merupakan cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data yang dimiliki.
- b. Estimasi
Estimasi hampir sama dengan klasifikasi, kecuali *variable* target estimasi lebih ke arah numerik daripada ke arah kategori. Model yang dibangun menggunakan *record* lengkap yang menyediakan nilai *variable* target sebagai nilai prediksi.
- c. Prediksi

Prediksi menerka sebuah nilai yang belum diketahui dan juga memperkirakan nilai untuk masa mendatang.

d. Klasifikasi

Dalam klasifikasi terdapat target *variable* kategori, misal penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu tinggi, sedang, dan rendah.

e. Pengklasteran

Merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

f. Asosiasi

Asosiasi bertugas menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

2.4. Algoritma FP-Growth

Frequent Pattern Growth (FP-Growth) adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data [2]. Algoritma FP-Growth merupakan pengembangan dari algoritma Apriori. Sehingga kekurangan dari algoritma Apriori diperbaiki oleh algoritma FP-Growth.

FP-Growth menggunakan konsep pembangunan tree dalam pencarian frequent itemsets. Hal tersebutlah yang menyebabkan algoritma FP-Growth lebih cepat dari algoritma Apriori. Karakteristik algoritma FP-Growth adalah struktur data yang digunakan adalah tree yang disebut dengan FP-Tree. Dengan menggunakan FP-Tree, algoritma FP-Growth dapat langsung mengekstrak frequent itemset dari FP-Tree.

FP-tree merupakan struktur penyimpanan data yang dimampatkan. FP-tree dibangun dengan memetakan setiap data transaksi ke dalam setiap lintasan tertentu dalam FP-tree. Karena dalam setiap transaksi yang dipetakan, mungkin ada transaksi yang memiliki item yang sama, maka lintasannya memungkinkan untuk saling menimpa. Semakin banyak data transaksi yang memiliki item yang sama, maka proses pemampatan dengan struktur data FP-tree semakin efektif.

Penggalian itemset yang frequent dengan menggunakan algoritma FP-Growth akan dilakukan dengan cara membangkitkan struktur data tree (FP-Tree). Metode FP-Growth dapat dibagi menjadi 3 tahapan utama yaitu sebagai berikut :

a. Tahap Pembangkitan conditional pattern base

Conditional Pattern Base merupakan subdatabase yang berisi *prefix path* (lintasan prefix) dan *suffix pattern* (pola akhiran). Pembangkitan *conditional pattern base* didapatkan melalui FP-tree yang telah dibangun sebelumnya.

b. Tahap pembangkitan conditional FP-Tree

Pada tahap ini, *support count* dari setiap item pada setiap *conditional pattern base* dijumlahkan, lalu setiap item yang memiliki jumlah support count lebih besar sama dengan minimum *support count* ξ akan dibangkitkan dengan *conditional* FP-tree.

c. Tahap pencarian frequent itemset

Apabila Conditional FP-tree merupakan lintasan tunggal (*single path*), maka didapatkan *frequent itemset* dengan melakukan kombinasi item untuk setiap

conditional FP-tree. Jika bukan lintasan tunggal, maka dilakukan pembangkitan FP-growth secara rekursif.

3. HASIL DAN PEMBAHASAN

3.1 Data mining dengan Algoritma FP-Growth

Dalam penelitian ini menerapkan *Data Mining* aturan asosiasi dengan algoritma *FP-Growth* dalam menganalisa data penjualan spare part di TB-Damar untuk mendapatkan pola pelanggaran yang terjadi.

Untuk melakukan penggalian dari data transaksi penjualan *spare part*, penulis menggunakan sebuah algoritma yang menjadi dasar dari algoritma-algoritma yang lain yaitu *frequent pattern growth (FP-Growth)*. Frekuensi kemunculan tiap *item* dari data transaksi dapat dilihat pada tabel dibawah ini :

Tabel 1 *Frequent pattern growth (FP-Growth)*

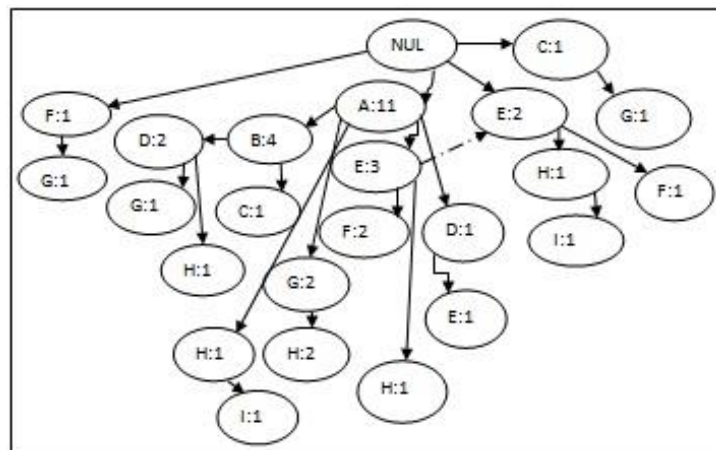
Trasn.	Itemset	Frek.
1	AXLE	11
2	BATTERY ASSY (12V 3.5A) 3C	3
3	BRAKE PAD KIT (54P2)	3
4	BUSH (3C11)	3
5	CABLE, CLUTCH (3C11)	6
6	GASKET, CYLINDER (KT) RXS	4
7	GRAPHIC 1 (UR YB 2006) SCORPIO-Z	5
8	OIL SEAL (45P1)	6
9	SCREW, VALVE ADJUSTING (3C11)	2

Tabel 2 Pembentukan FP-Tree

Trans.	Dataset Penjualan Spare Part
1	{ a, e, f }
2	{ e, h, i }
3	{ a, b, c }
4	{ a, d, b, g }
5	{ d, e, a }
6	{ c, g }
7	{ a, b }
8	{ e, f }
9	{ f, e, a }
10	{ d, b, a, h }
11	{ a, g, h }
12	{ f, g }
13	{ e, a, h }

Trans.	Dataset Penjualan Spare Part
14	{ a, h, g }
15	{ i, a, h }

Dari Tabel 2 diatas, langkah selanjutnya yang harus dilakukan adalah membentuk jalur FP-Tree. Dimana hasil pembentukan jalur 1 sampai 15 dapat dilihat di gambar berikut :



Gambar 1. Jalur FP-Tree

3.2. Penerapan FP-Growth

Untuk menerapkan algoritma *FP-Growth* pada data transaksi penjualan spare part maka tahapan yang harus dilakukan untuk mencari *frequent itemset* adalah membuat *FP-Tree*. Setelah *FP-Tree* terbentuk, maka langkah selanjutnya adalah tahap menentukan *Conditional Pattern Base*, *Conditional FP-Tree*, dan *Frequent Itemset*. ketiga tahapan ini dapat dibuat dengan melihat kembali *FP-Tree* yang sudah dibuat sebelumnya.

Hasil dari pembangkitan *FP-Tree* maka didapatkan *Conditional Pattern Base*. Berikut ini hasilnya :

Tabel 3 *Conditional Pattern Base*

Suffix	Conditional Pattern Base
i	{ (e,a:1), (a,h:1) }
h	{ (e:1), (a,b,d:4), (a,e:1) }
g	{(a,b,d:1), (a:2), (f:1), (c:1)}
f	{ (a,e:2), (e:1)}
e	{ (a:3),(d:1) }

Suffix	Conditional Pattern Base
d	{(a,b:2)}
c	{(a,e:2)}
b	{(a:4)}
a	

Untuk mencari *Conditional FP-Tree* adalah dengan cara menjumlahkan *support count yang ada* dan *support count yang lebih besar* akan dibangkitkan dengan *conditional FP-tree*. Berikut tabel hasil *Conditional FP-Tree* :

Tabel 4 *Conditional FP-Tree*

Conditional FP-Tree
{ (e,a:1), (a,h:1) }
{ (e:1), (a,b,d:4), (a,e:1) }
{(a,b,d:1), (a:2), (f:1), (c:1)}
{ (a,e:2), (e:1)}
{ (a:3),(d:1) }
{(a,b:2)}
{(a,e:2)}
{(a:4)}

Setelah mencari *Conditional FP-tree* maka tahapan yang selanjutnya adalah mencari *Frequent Itemset*. Dimana tahapan ini mencari *single path* kemudian dikombinasikan dengan *item* yang ada pada *Conditional FP-Tree*. Tabel berikut ini hasil dari *Frequent Itemset* :

Tabel 5 *Frequent Item Set*

Frequent Itemset
(h,i:2)
(e,h):2,(a,h):5,(b,h):4,(d,h):4
(a,g):3
(a,f):2,(e,f):3
(a,e):3
(a,d):2,(b,d):2

Frequent Itemset

(a,c):2,(b,c):2

(a,b):4

Jika *minimum support* 40% dan *minimum confidence* 60%, maka *strong association rule*-nya dapat dilihat pada table berikut ini :

Tabel 6 *Strong Association Rule*

Item A	Item B	Support	Confidence
Screw valve adjusting	Oil seal	13.33%	100%
Oil seal	Axle	33.3%	83.3%
Graphic 1	Axle	20%	60%
Gasket cylinder	Cable cluth	20%	75%
Brake	Axle	13.33%	67%
Bush	Axle	13.33%	67%
Axle, cable cluth	Oil seal	13.33%	67%
Cable cluth, gasket cilynder	Gasket cylinder	13.33%	67%
Battery assy	Axle	20%	100%
Oil seal, graphic 1	Axle	13.33%	100%
Axle, gasket cylinder	Cable cluth	13.33%	100%

3.3. Implementasi dan Pengujian

Implementasi *Data Mining* dalam menganalisa transaksi penjualan spare part ini menggunakan aturan asosiasi dengan algoritma *FP-Growth*. Algoritma *FP-Growth* sebagai proses utama dari *association rules* dimulai dengan membentuk kandidat *itemset*, membentuk *FP-Tree*, kemudian menentukan *Conditional Pattern Base*, *Conditional FP-Tree* dan *Frequent Itemset*. Sebelum melakukan pengujian terhadap data, maka yang dilakukan adalah mempersiapkan data transaksi penjualan spare art yang berjumlah 150 kemudian data diinputkan ke dalam *Microsoft Excel*. Dimana hasil pengujian *Association Rule Text View* yang dihasilkan oleh aplikasi *RapidMiner* dapat dilihat pada gambar.


```
AssociationRules  
  
Association Rules  
[graphic 1] --> [axle] (confidence: 0.600)  
[bush] --> [axle] (confidence: 0.667)  
[brake] --> [axle] (confidence: 0.667)  
[axle, graphic 1] --> [oil seal] (confidence: 0.667)  
[axle, cable clutch] --> [gasket cylinder] (confidence: 0.667)  
[cable clutch, gasket cylinder] --> [axle] (confidence: 0.667)  
[gasket cylinder] --> [cable clutch] (confidence: 0.750)  
[oil seal] --> [axle] (confidence: 0.833)  
[battery assy] --> [axle] (confidence: 1.000)  
[screw valve adjusting] --> [oil seal] (confidence: 1.000)  
[oil seal, graphic 1] --> [axle] (confidence: 1.000)  
[axle, gasket cylinder] --> [cable clutch] (confidence: 1.000)
```

Gambar 2. Hasil *Association Rules*

Maka hasil dari perhitungan menggunakan *RapidMiner* menghasilkan 12 *rules*. Hasil yang didapat dari *Association Rules (Text View)* juga menjelaskan bahwa item a dan b saling berkaitan dengan menjelaskan secara *text*. Hasil perhitungan diatas berbeda dengan perhitungan data sampel, Karena jumlah data yang diolah ini berjumlah 150 data lebih banyak dari sampel data yang ada berjumlah 15 data.

4. KESIMPULAN DAN SARAN

Dari hasil penelitian yang dilakukan, maka kesimpulan yang dapat diambil berdasarkan pembahasan implementasi *Data Mining* dalam penjualan *spare part* motor *sport* dengan menggunakan algoritma *FP-Growth* adalah :

- a. Hasil penjualan dari *spare part* motor *sport* yang paling banyak terjual di TB-Damar CV.TJAJAJA BARU Padang bisa diketahui dengan menggunakan algoritma *FP-Growth*. *Spare part* yang memenuhi *minimum support* dan *minimum confidence* serta yang banyak terjual adalah *screw valve adjusting*, *oil seal*, *battery assy*, *axle*, *gasket cylinder*, dan *cable clutch*.
- b. Menerapkan metode *Data Mining* dengan algoritma *FP-Growth* kedalam aplikasi untuk analisis pola pembelian konsumen sangat bermanfaat bagi perusahaan tersebut, karena TB-Damar akan mengetahui *spare part* mana yang banyak dibeli dan membantu dalam pemesanan *spare part* pada kantor pusat.

Implementasi algoritma *FP-Growth* pada aplikasi *RapidMiner* dimulai dari input data penjualan yang akan menjadi *database* pada *Ms.Excel*, kemudian data akan diproses menggunakan aplikasi *RapidMiner*, lalu pembentukan *support* dan *confidence*, dan terakhir akan menghasilkan *strong rule*.

Dari hasil penelitian ini, penulis dapat memberikan beberapa saran yang dapat diperhatikan untuk penelitian dalam pengembangan selanjutnya, yaitu :

- a. Dalam implementasi data ini, jika semakin besar data yang diambil, maka akan mendapatkan hasil yang lebih akurat dan lebih besar juga. Jika dilakukan pengembangan terhadap penelitian ini, sebaiknya dilakukan penambahan *sample* juga.

- b. Salah satu teknik yang digunakan dalam penelitian ini yaitu *RapidMiner*, selain itu bisa juga menggunakan aplikasi *Data Mining* lainnya seperti *Tanagra* atau *WEKA*.
- c. Dalam pembagian *spare part* harus sesuai dengan merk dan jenis, agar konsumen lebih jelas dan akurat dalam pemesanan atau penjualan *spare part* motor, hal ini berguna untuk memudahkan konsumen dalam pembelian dan admin dalam pendataan transaksi penjualan tersebut.

DAFTAR PUSTAKA

- [1] Andriani, Anik. 2013. ***Aplikasi Data Mining Market Basket Analysis Penjualan Suku Cadang Sepeda Motor Menggunakan Metode Association Rules Pada PT.Sejahtera Motor Gemilang.*** Jurnal Ilmiah Teknik Informatika Universitas Nusantara PGRI.Kediri.
- [2] Sijabat, Alimancon. 2015. ***Penerapan Data Mining Untuk Pengolahan Data Siswa Dengan Menggunakan Metode Decision Tree.*** Jurnal Ilmiah Teknologi dan Informasi ASIA Vol. 8 No 2, Agustus 2014.
- [3] Fadlina. 2014. ***Data Mining Untuk Analisa Tingkat Kejahatan Jalanan Dengan Algoritma Association Rule Metode Apriori.*** Jurnal Ilmiah Informasi dan Teknologi Ilmiah 2014,Padang.
- [4] Santoso, Leo. 2014. ***Pembuatan Perangkat Lunak Data Mining Untuk Penggalan Kaidah Asosiasi Menggunakan Metode Apriori.*** Teknik Informatika Universitas Kristen Petra.
- [5] Erwin. 2009. ***Analisis Market Basket dengan Algoritma Apriori dan FP-Growth.*** Jurnal Generic 26-30.
- [6] Tyas Eko D. 2008. ***Penerapan Metode Association Rules Menggunakan Algoritma Apriori Untuk Analisa Pola Data Penjualan Motor.*** Jurnal ilmiah Bisnis & Teknologi ISSN : 2407-6171 SEMBISTEK 2008 Jakarta.
- [7] Sijabat, Alimancon. 2015. ***Penerapan Data Mining untuk Pengolahan Data Siswa dengan Menggunakan Metode Decision Tree.*** Jurnal Informasi dan Teknologi Ilmiah. Volume 5 No 3. ISSN : 2339-210X.
- [8] Sunjana. 2010. ***Klasifikasi Data Nasabah Sebuah Asuransi Menggunakan Algoritma FP-Growth.*** Seminar Nasional Aplikasi Teknologi Informasi 2010 (SNATI 2010) ISSN: 1907-5022.
- [9] Tampubolon Kennedi, dkk. 2013. ***Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat - Alat Kesehatan.*** Informasi dan Teknologi Ilmiah(INTI).Volume 1, Nomor 1, Oktober 2013. ISSN : 2339-210X.
- [10] Kusri dan Emma Taufiq Lutfi. (2009), ***Algoritma Data Mining.*** Andi,Yogyakarta.
- [11] Algifanri Maulana. (2016). ***The Influence Of Dokuku Data Security and Service-Oriented Architecture On The Quality of Information.*** *Teknologi Dan Sistem Informasi (TEKNOSI)*, Vol 2 No 3(Universitas Andalas), 41-46.
<https://doi.org/92-502-1-PB>