MODERN APPLIED SCIENCE

Vol. 6, No. 10 October 2012

CANADIAN CENTER OF SCIENCE AND EDUCATION

Editorial Board

Contents

Fatigue Behaviour of Steel Bridge Joints Strenghtened with FRP Laminates

Alessio Pipinato[1], Carlo Pellegrino[1] & Claudio Modena[1]

[1] University of Padova, Padova, Italy

Correspondence: Alessio Pipinato, Department of Civil and Environmental Engineering, via Marzolo 9, Padova 35131, Italy. Tel: 39-42-533-292. E-mail: alessio.pipinato@unipd.it

Abstract

One of the most relevant concerns about steel and metal bridges stands on the repair and rehabilitation of existing structures. In fact, the remaining service life of steel bridges is limited by fatigue damage and in order to ensure the safety of these bridges it is often necessary to inspect the structure in order to discover the presence of fatigue cracks. New reinforcement techniques are needed, in order to prevent fatigue cracking and to increase bridge safety at the same time. According to the contemporary knowledge on the matter, two alternatives of rehabilitation are given: the traditional interventions and the use of carbon fibre composites. In this paper, in order to explain the appropriate use of these techniques, the most common knowledge on the matter are presented and explained, focusing in particular on the real scale testing results coming from the last decades research applications related to steel bridge engineering. This study should in this way be useful for future research applications and for bridge engineering real case rehabilitations.

Keywords: fatigue, steel, bridge, rehabilitation, FRP

1. Introduction

The remaining service life of steel bridges is limited by fatigue damage and in order to ensure the safety of these bridges it is often necessary to inspect the structure in order to preclude the presence of fatigue cracks. Damage due to cyclic loadings deals with a fatigue failure that proceeds by gradual enlargement of fatigue cracks during the tensile part of the loading cycle; once a crack reaches a certain critical size, catastrophic failure ensues. The process of fatigue crack growth is evidenced by the formation of characteristic striation patterns on the fracture surface called beach markings, which are followed by the more irregular pattern characteristic of fast fracture. In order to maintain these bridges in service new reinforcement techniques are needed, in order to prevent fatigue cracking and increase bridge safety at the same time. The use of carbon fiber composites (CFRP) for structural reinforcements has shown increasing promise. Aside from the very high tensile resistance, composite materials have also demonstrated a very high fatigue resistance, which makes them particularly appropriate for the reinforcement of steel structures subjected to fatigue loads. Degradation of steel structures, such as old industrial buildings and bridges, and increased load requirements have led to the need for structural rehabilitation and strengthening. Another rehabilitation procedure stands on traditional applications as HS-bolting or welded steel plates to the original structure, sometimes used as a strengthening technique: some negative aspects such as the increase in permanent loads, difficulty of applications, and problems due to corrosion and fatigue have been demonstrated to affect also the new members or connections. On the contrary, fiber reinforced polymer (FRP) materials have a high strength-weight ratio, do not give rise to problems due to corrosion, are extremely manageable and are commonly used for strengthening structural existing members (Lua et al., 2005; Yuana et al., 2004; Jao et al., 2005; Pellegrino et al., 2002), but they have only been infrequently studied and used for strengthening steel elements and, above all, steel-concrete composite elements. Several studies on the FRP strengthening of concrete structures have resulted in the first design guidelines for concrete structures strengthened with externally applied FRP. American ACI 440-02 (ACI, 2002), European fib bulletin 14 (FIRB, 2001) and Italian Recommendations (CNR, 2004) are examples of such guidelines. Although externally bonded FRP strengthening for metal structures is a rapidly developing technique (Cadei et al., 2004), guidelines for the reinforcement of steel structures with FRP do not currently have a comparable degree of reliability. Some examples of guidelines for the design and construction of externally bonded FRP systems for strengthening

existing metal structures are the ICE design and practice guide (Moj, 2001), CIRIA Design Guide (CIRIA Design Guide), US Design Guide (Schnerch et al., 2007) and document CNR-DT 202/2005 (CNR, 2005), which can currently be considered as a preliminary study. The benefits of composite strengthening are shown in some significant case studies like the successful strengthening of a steel bridge on the London Underground (Moj, 2007). The benefits of strengthening large cast-iron struts with Carbon FRP composites in the London Underground are illustrated in (Moj, 2006). Significant work in this area of research was also developed by (Tavakkolizadeh et al., 2003a; Tavakkolizadeh et al., 2003b; Shaat et al., 2006; Shaat & Fam, 2004; Sen et al., 1995; Miller et al., 2001). A state-of-the-art review on FRP strengthened steel structures was recently developed by Zhao and Zheng (2007). Main approaches proposed in the literature to study the phenomenon of delamination (Buyukozturk et al., 2004) are based on linear elastic analysis of internal stresses (Taljsten, 1997) and linear elastic fracture mechanics (Lenwari, 2005; Lenwari et al., 2006; Colombi, 2006). Other significant approaches to the determination of stresses in the adhesive layer are described in the CIRIA Design Guide (Cadei et al., 2004) and in the work of Deng et al. (2004). The fundamental work about stresses at the interface was reported in Smith and Teng (2001) and in Stratford and Cadei (2006). These theories have not been validated with a sufficient number of experimental results, due to the few experimental data available in the literature. One of the first studies on the flexural behaviour of steel beams reinforced with carbon-fibre reinforced polymer composites is that of Moy and Nikoukar (2002). Very recent works about strengthening steel structures using FRP are those of Schnerch and Rizkalla (2008) and Rizkalla et al. (2008). Current design guide lines propose a method to predict the flexural behaviour of FRP strengthened elements based on classic equilibrium and strain compatibility, usually within the elastic range of the materials. FRP laminate is always considered only under the tensile flange current in-depth information about the non-linear behaviour of FRP strengthened steel-concrete composite elements is scanty. Concerning the phenomenon of fatigue in existing steel bridge member, some works could be noticed in literature, dealing in particular with the assessment of riveted connections, both by full scale tests, or by step level procedure, or finally applied on real case applications (Pipinato et al., 2012a; Pipinato et al., 2012b; Pipinato et al., 2012c; Pipinato et al., 2012d; Pipinato et al., 2011a; Pellegrino et al., 2011; Pipinato, 2011; Pipinato et al., 2011b; Pipinato et al., 2010; Pipinato & Modena, 2010; Pipinato, 2010; Ppinato et al., 2009; Pipinato et al., 2008). This paper deals with the analysis of experimental tests developed in a variety of steel joints repaired with FRP laminates. The purpose of the study is to investigate three main groups of test: (a) tension tests on single plates with FRP reinforcement; (b) tension test on double plates with FRP reinforcement; (c) bending tests on reinforced full scale beams with FRP reinforcements on the lower flange. SN curves are provided in order to check the correspondence of experimental results and coded provisions.

2. Existing Metal Members

Iron bridges have been built since the industrialisation time period, which was the end of the 19th century, so old metal bridges could imply to assess a structure with an average age of more than 100 years. Several bridges even exist since the mid of the nineteenth century, leading to an age of more than 150 years. The early metal bridges in the 19th century were fabricated with cast iron or puddle iron (wrought iron). The manufacturing of puddle iron started in the beginning of the 19th century. Since it had a lower carbon content than cast iron, going along with a better ductility, it allowed forging and an easier workmanship. Yet towards the end of the 19th century puddle iron was superseded by mild steel that obtained higher qualities concerning the chemical analysis and cleanness of the steels as well as better technological material properties (e.g. weldability, strength). The steel production and construction technology (bolting and welding of joints) developed quickly but testing methods to examine relevant properties as toughness, fatigue, corrosion etc., were missing completely. Such testing methods were developed much later in the 20th century and related to modern steels rather than old steels of existing structures. Therefore only fragmentary knowledge exists concerning iron materials of the early times, complicating the handling and assessment of old metal structures. Already at the beginning of the 20th century metal bridges were built mainly with mild steel. It should be also considered that before 1910, there was little standardization in the industry: each steel producer used his own recipe and rules. This resulted in a wide variety of metals used, regarding the chemical and mechanical properties. Railway authorities, for instance, adopted specific rules only from 1900 in order to specify the materials to be used for iron railway bridges. In order to assess existing bridges it is essential to know the material properties and their characteristics. Constitutive materials could be: grey cast iron, wrought iron, mild steel and the contemporary steel product. Grey cast iron takes its name from the colour of the fracture face, has a carbon content of 1.5-4.3% and 0.3-5% silicon plus manganese, sulphur and phosphorus. It is brittle with low tensile strength, but is easy to cast. The characteristics of the material are good in compression but poor in tension. Thus, the structural parts where often designed to be in compression as arches, columns etc. The graphite flakes caused a significant brittleness of the material. Internal cracks could easily occur and propagate along the flakes when the iron was subjected to tensile stresses,

giving poor ductility properties. This material has a good wear resistance and damping abilities, absorbing vibrations and noise, even if it is brittle and has demonstrated a poor resistance to impact and shock. Cast iron is not suitable for welding due to its high carbon contents, which can lead to brittle cracks in and around welded joints. Wrought iron (or puddle steel) was the first structural steel until it was replaced by (mild) steel, in the end of the 19th century. It has a low carbon content, an high amounts of phosphor and nitrogen making the material brittle and escalating the ageing process. The microstructure is non homogenous due to the manufacturing process producing inclusions of sulphides and oxides. This led to anisotropy of the material which is especially bad in the thickness direction due to the arrangement of the inclusions and the influence of the rolling. As a matter of fact, the mass production of steel started with the Bessemer process 1856, followed by the Martin-Siemens process 1867 and the Thomas-Gilchrist process 1878. Most of the old metal bridges still used today consisted of steel produced with one of these processes. Production of steel in the end of the 19th and the beginning of the 20th century were conducted with a technique called chill module casting. The chill module casting was performed by pouring the steel from the oven in to a chill module to cool down before rolling of the steel. The cooling process in the module started from the borders, with high temperatures in the middle. During the cooling process almost pure steel formed at the borders and unwanted alloys and impurities increases towards the centre of the melt (depending on thermal properties). This manufacturing process lead to the formation of impurities and blisters increased in the middle of the steel, and for this principle reason steel produced during these circumstances are not considered appropriate as construction steel today. Moreover, an high concentrations of unwanted compounds formed in the middle of the steel that drastically lower the quality, so laminates manufactured with this technique are characterized by steel with very good qualities at the surface while the centre of the plate will be steel with brittle properties. At the same time, they are not suitable for welding, due to a differentiated toughness, and for this reason cracks can originate due to the residual stresses from the heat affected zone of the weld. Finally, contemporary steel materials, are coded and produced according to general standards and methodologies, for e.g. in Europe according to EN 1993-1-1 (2005), characterized by different material properties as the tensile resistance, the fracture toughness, the ductility requirements.

3. FRP Materials

The principal fibres that are used in FRP plate bonding to metallic structures are carbon fibres, aramid fibres and E-glass fibres: a summary of the indicative fibre properties is given in Table 1. Carbon fibre is characterized by excellent fatigue resistance, and do not suffer from stress rupture compared with glass or aramid fibres. Carbon fibre properties depend on the structure of the carbon used, so typically they come defined as standard, intermediate and high modulus fibres, depending from their structural behaviour. Aramid fibres have the highest strength to weight ratio (for e.g. Kevlar) and is characterized by similar tensile strength to glass fibre, but can have modulus greater at least two times; it could allow significant energy absorption but, compared to carbon, it is lower in compressive strength and has poorer adhesion to the matrix, finally it is also susceptible to moisture absorption. E-glass accounts for the larger part of the glass fibre use and is used mainly in a polyester matrix: the 'E' in E-glass stands for 'electrical' and was intended to indicate that the material has low electrical conductivity. It is commonly the the cheapest glass fibre and is therefore the preferred material in the construction market. Two types are available one contains boron, the other boron free: the corrosion resistance of E-glass without boron is approximately seven times the corrosion resistance of the boron-containing E-glasses.

4. Experimental Tests

The first group of experimental tests investigated, is represented by FRP reinforced steel plates under tension forces (Figure 1) (Jones & Civjan, 2003). Examined details were notched (N) or hole (H) provided: N11 not reinforced; N21 is made of 4 Sika Wrapex 103C resin on one side; N31 is made of Sika Wrapex 103C resin on two sides, L=255 mm; N41 is made as N31 but with a smoothed resin layer; N35 is made as N31 but with an unbalanced resin mix; N51 is made as N31 but with L=380 mm; N22 is made as N21 but with MBRACE CF130; N36 is made as N31 is made as N21 but with MBRACE CF130; H11 is not reinforced; H21-22 is made of Sika Wraplex 103 C on two sides, L=255 mm; H41-42 as H21 but with Sika Carbodur. In this test the FRP application caused eccentricity causing bars curvature, and as a result the performance doesn't grow up if compared to the unreinforced test. Three types of CFRP have been used with diverse elasticity module, and varying thickness: from these tests it could be obtained that the best performance is obtained by using the CFRP with a lower E module (e.g. Sika WrapHex 103C). Moreover, from the obtained comparison on SN curves, an extended life has been obtained for: holed specimens, accurate preparation of the base layer, longer CFRP laminates. Category details, even if approximate, stands in the range C=140-160 for notched specimens, and on C=160 for holed specimens. A similar experimental testing has been provided as reported in Figure 2 (Monfared et al., 2008): in this test, similar specimens of the aforementioned test has been used. Variable parameters have

been used, as: a reinforcement on one or two sides, and the result was that a single side specimen has a longer fatigue life; an accurate sanded surface preparation, belonging to the result that the surface preparation doesn't provide significant improvements. While in the reinforced specimens the failure stands on the reinforment, in the not-sanded specimens the CFRP delaminates. In Figure 2 it could be noticed a tendency line for the failure specimens, belonging to the following detail category: C=78 for not-reinforced specimens, and C=82 for reinforced specimens. Another similar application has been developed (Zheng & Lu, 2006) and reported in Figure 3: in this study the reinforced detail on both sides highlighted a longer fatigue life, which is improved also by higher stiffness value, depending on the E module and on the thickness. The detail failed for delamination. The following detail categories were obtained: C=63 for not-reinforced specimens, and C=102 for reinforced specimens. The second group of experimental investigations analyzed, is represented by double strap FRP reinforced steel plates under tension forces (Figure 4) in this case the failure mechanism is represented by delamination for SDS specimens, and CFRP failure for the DDS, beyond the C=40 (Matta & Karbhari, 2005); in the second case (Liu et al., 2005), the failure mechanism is influenced by the CFRP module; with a normal module CFRP delamination prevail, while with higher module CFRP the composite laminates fails. Other studies, presents some analysis on existing bridge steel members specimens of the first of '900, taken out of service and repaired with CFRP: in the first cases (Bassetti, 2001) the CFRP stiffness is varying from SikaCarboDur S512 with a lower module (E=155 GPa) and laminates of 1.2 mm thick, to higher module (E=210 GPa) and laminates of 1.4 mm thick; the mechanical strength resulted to be not influenced from the quality of the resin, while the pre-tension of the CFRP could grow-up to 20 times the fatigue life; failure resulted to remain in the delamination of the CFRP. Other observations has been done in (Täljsten et al., 2001) in relation to the propagation speed growth of the CFRP, which has been demonstrated to depend from the pretension force, from the stress reversals and pre-tension, and from the importance of the cleaningness of the surface; as a result, pre-tensioned specimens could stop the crack propagation into the very high cycle phase: for e.g. a pre-tension specimen has been tested up to 16x106 cycles, belonging to an extended life raising eight times the category detail life. Moreover, the crack propagation has been observed to be symmetric except if a delamination raise up. In the graph reported in Figure 6 two groups has been presented: in red not-reinforced details, in blue reinforced details, with the further specifications denoted in the legenda. From these comparisons, it could be noticed some SN tendencies, such as C=100 for virgin specimens, C=106 for not reinforced beams, and C=152 for reinforced beams. The third group of experimental investigations analyzed, is represented by bending tests on steel beam (Figure 7): four riveted beams have been tested by bending fatigue tests. Reinforcing technique are for e.g. represented by cover-plating or rivet/bolt substitution or by stop-holes: the limits observed in the testing results are that these are applicable for small cracks, and could be applied in the first propagation phase; moreover, these technique are limited to local repairing. CFRP interventions are otherwise reported in the following Figure 8, in which not reinforced and a corresponding CFRP reinforced section has been tested; in particular for the Q4 specimen, it should be observed that the test has been stopped at 20 milion of cycles without failures nor crack propagation, while in the Q5 and Q8 specimens, even if large crack openings has been found before testing, these has been limited in the growing phase, and no CFRP reinforcement has been found at the end of the test. To sum up all the tests analyzed, the principles data have been reported in Tables 2 and 3.

Table 1. Typical fiber properties used in structural interventions

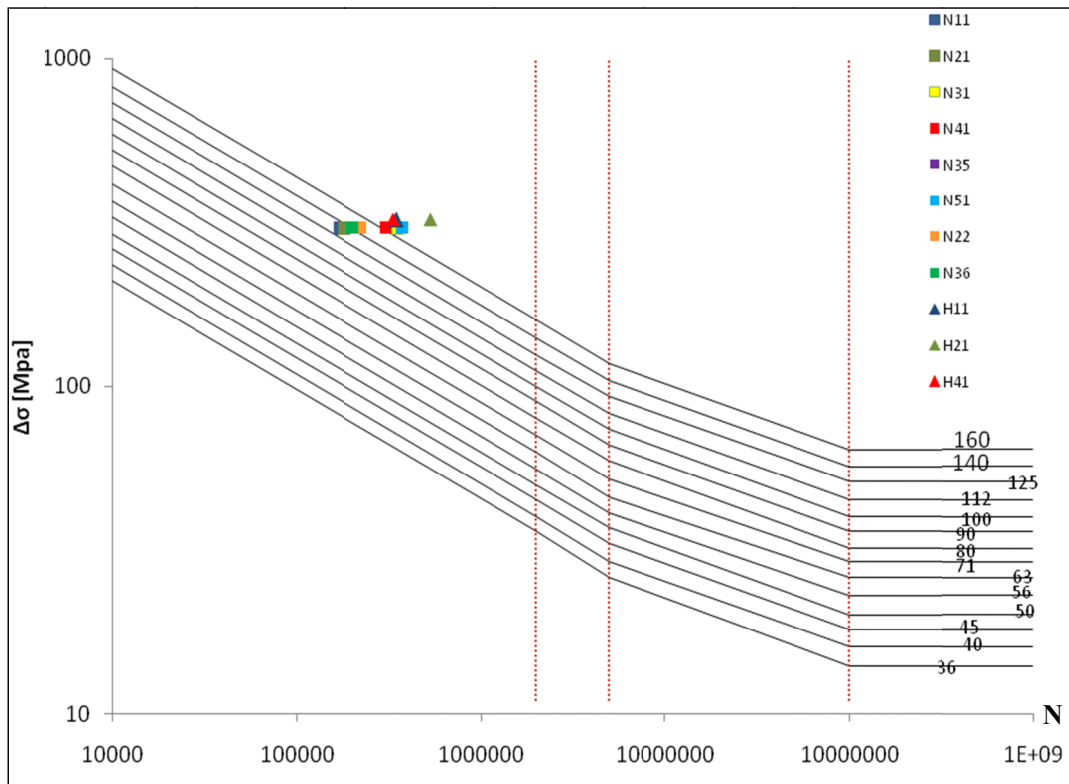| Property	Ultra-high-modulus carbon fiber	High strength carbon fiber	High-modulus carbon fiber	Aramid fiber Kevlar 49	E-glass fibre
Specific weight	2.12	1.80	1.80	45	2.56
Modulus of elasticity (GPa)	620-935	230	390	130	70
Tensile strength (MPa)	3600-3700	3400	2900	3000	-
Strain to failure (%)	0.6	1.48	0.74	2.3	-
Coefficient of thermal expansion (10^{-6} °C)	-	-1.0 to +0.4	-1.0 to +0.4	-5.2	4.9

Figure 1. Fatigue failures of FRP reinforced steel plates subjected to tension forces
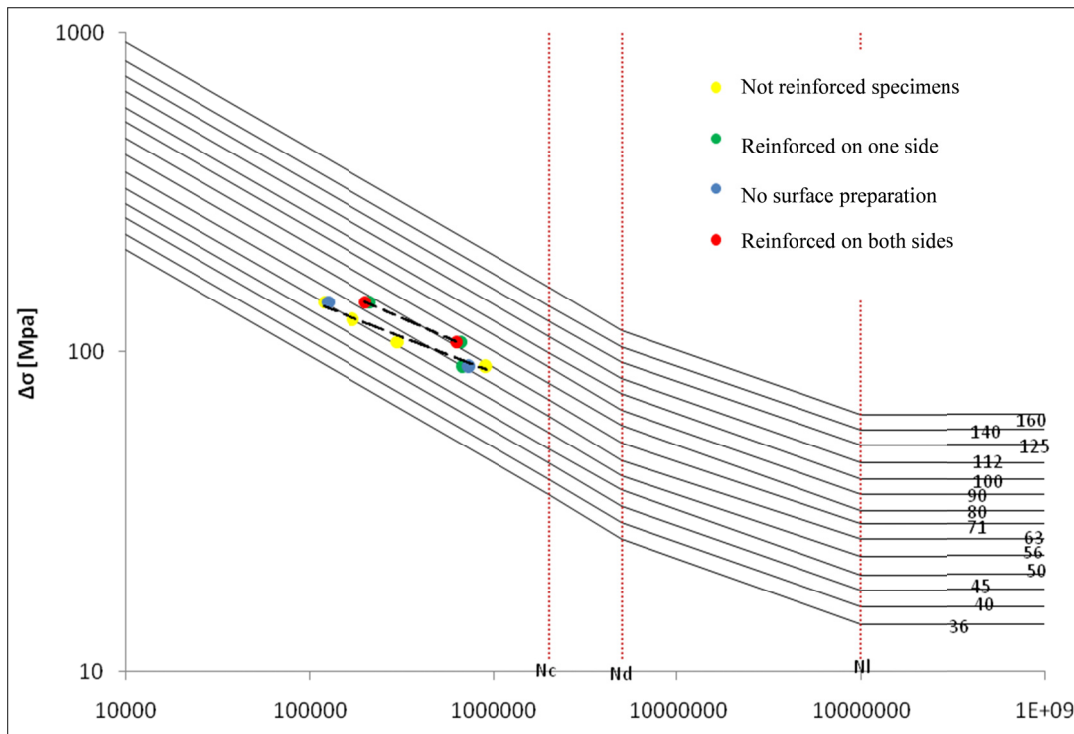

Figure 2. Fatigue failures and curve of FRP reinforced steel plates subjected to tension forces

Table 2. Bending test, FRP reinforced

	Riferimento	Campioni	Tipo di test	Dim. materiali	Caratteristiche materiali	Dettaglio	Δσ [Mpa]	N	Note sulla rottura	Immagine esemplificativa
10	[Tavakkolizadeh 2003]	5 steel beams A36 notched on flanges and reinforced with CFRP on the lower side at the midspan	Bending test	Steel beam: S127x4.5, 1.22m long. CFRP: 300mmx76 mmx1.27 mm.	Fy=336.4 Mpa, 330.9 Mpa, E=194.4 GPa .CFRP: E=144 Gpa,shear strenght=2137 Mpa	Beam reinforced at the lower flange at midspan	207 241 276 310 345	1.191.140 71.278 35.710 30.216 19.068	Crack initiation and growth are similar in reinforced specimens, but quite different in the not-reinforced specimen. Crack initiate always in the edge of notches going up to the beam panel.	
11	[Deng 2007]	UB13 steel beams CFRP reinforced on the lower flanges at midspan	Bending test	Steel beams: 127x76UB13, 1.2 m long. FRP: 3 mm thick, 400 mm long.	Steel E=205 Gpa. Sikadur 31 Normal, E = 8 Gpa, ft= 29.7 Mpa.	Beam reinforced at the lower flange at midspan	39 30 27	20.000 130.000 965.000	The crack propagates from the mid of the adhesive up to the steel-adhesive interface, with an angle of 45°.	
12	[Bassetti 1999]	Variable sections steel riveted beams reinforced with traditional tecniques.	Bending test	Beam: 5700mm long, angle profiles 100x100x10 mm, plates: 230x10 mm	Steel beam: E=212 Mpa, fu=412 Mpa, fy=312 Mpa.	QT7 one plate reinforced QT6 rivets replaced with HS bolts QT6 with stop holes QT6 with stop holes and HS bolt QT2 rivets replaced with bolt QT2 rivets replaced with bolt and a stop hole QT1 rivets replaced with bolt	80 72 91 72 71 71 91	5.918.000 12.177.000 16.644.000 17.185.000 6.217.000 7.391.000 6.144.000	In these beams the failures stands commonly on rivet holes. In case of short cracks, propagation could be stopped replacing rivets with HS bolts.	
13	[Bassetti 2001]	Variable sections steel riveted beams not reinforced.	Bending test	Beam: 5700mm long, angle profiles 100x100x10 mm, plates 230x10 mm.	Steel beam: E=212 Mpa, fu=412 Mpa, fy=312 Mpa.	QT4	80	3.532.000	The failure of the beam started from the rivet holes.	
						QT5	80	2.109.000	The failure of the beam started from the rivet holes.	
						QT8	80	7.653.000	The failure of the beam started from the rivet holes.	
		Variable sections steel riveted beam, with 2 stripes of CFRP normal strenghtening and 3 stripes of CFRP post-tensioned strenghtening	Bending test	Beam: 5700mm,angle profiles 100x100x10 mm, 230x10 mm plates. CFRP thickness: 1.4 mm	Steel beam: E=212 Mpa, fu=412 Mpa, fy=312 Mpa. CFRP SikaCarboDur M614.	QT4	80	20.229.000	QT4: the test has been stopped with no failures; QT5, QT8: crack propagation observed.	
						QT5	80	10.440.000		
						QT8	80	2.392.000		

Table 3. Tensile test on single steel plates, FRP reinforced

	Reference	Specimens	Test type	Dimensions	Materials	Details	Δσ [Mpa]	N	Failure notes	Image
1	[Bocciarelli 2009]	S275 specimens reinforced on both side with Sika carbodur M614	Tensile test	Steel specimens: 1000 mm long, 50 mm large, 6 mm thick. FRP:500x50x1.4mm.	Steel S275: E=21 GPa, nominal strenght 331 Mpa, minimal tension strenght 430 MPa.CFRP SikaCarboDur M614: E > 200 Gpa, ft > 2800 Mpa. Sikadur resin 30: E = 12840 Mpa, ft=30.2 Mpa	FT1 FT2 FT3 FT4 FT5 FT6 FT7	83 83 100 100 120 120 160	5.800.000 4.900.000 1.800.000 1.700.000 1.200.000 1.200.000 250.000	Failure mechanism started with the detachment from the edge of the element and then propagates on the whole element interface. A visual inspection reveals that the adhesive-metal interface is the weakest point of the reinforced system.	
2	[Iwashita 2007]	Single-lap shear steel plates	Tensile test	Steel specimens: 400 mm x 12.5mm x 1.6mm. FRP: nominal thickness=0.128mm	Steel SM490: tension strenght 490 Mpa, E=210 Gpa. FRP: tension strenght 3500 Mpa, E=235 Gpa. Epoxy resin: tension strenght 51.9 Mpa, E=3.53 Gpa	SF7 SF71 SF66 SF55 SF44	195 179 162 130 98	4.412 6.255 14.058 34.362 93.199	The most common failure mechanism has been found to be the detachment of the resin from the steel plates.	
3	[Jones 2003]	5 steel plates not reinforced and 24 steel plates reinforced on one or both side	Tensile test	Steel specimens: 6.5x51x510mm. SikaCarboDur thickness: 1.2mm; SikaWrapHex thickness 103C:1.0 mm; MBRACE CF130 thickness:1.0 mm	Steel fy=340 Mpa, fu=483MPa. E(SikaCarboDur)=165000MPa; E(SikaWrapHex 103C)=65000MPa; E(MBRACE CF130)=38000MPa	N11 not reinforced N21: 4 Sika Wrapex 103C layer N31-N34: Sika Wraphex 103C on two sides L=255mm N22: as N21 but with Mbrace CF 130 N36: as N31 but with Mbrace CF 130 H21-H22: Sika Wraphex 103C on both sides L=255mm H41-42: as H21 but with Sika Carpodur	305 305 305 305 305 323 323	173.418 182.829 353.328 222.731 200.851 533.772 336.455	Failure has been observed for CFRP delamination for all tests, no failure has been observed in the fibers, fatigue cracks propagates bejond CFRP fibers.	
4	[Monfared 2008]	15 steel plates notchedm with and withou t FRP (on both sides), sanded or not.	Tensile test	Steel specimen: 9.62mm x 50.8mm x 350 mm. Dim. FRP: 1mm x 38mm x 250 mm.	Steel: nominal strenght =350 Mpa, E = 200GPa. E(CFRP)=607GPa, tension failure CFRP =2160MPa. E(adhesive)=12.8GPa	UR 1 UR-2 UR-3 UR-4 R1S-S-1 R1S-S-2 R1S-S-4 R2S-S-2 R2S-S-3 R2S-S-4	90 108 126 144 90 108 144 108 126 144	907.533 302.500 172.325 118.192 672.584 663.784 211.336 625.000 33.491 199.251	Failure lies on the FRP for reinforced specimen , for delamination in not sanded specimens.	
5	[Zheng 2006]	6 steel specimens hollowed at the centre, with 2 indentation, reinforced with CFRP on one or both side.	Tensile test	Steel specimens: 100x100x7mm. FRP: 100x300x1.0 mm, 100x300x1.4 mm.	Steel: fy=435 Mpa. FRP; En=165 Gpa, Eh = 320 Gpa. Aradilte adhesive shear E=2015 0.9 Gpa	-	120 90	1.000.000 3.312.000	The remaining fatigue life is extended from the 155 to the 580% for reinforced specimens. The propagation speed of cracks in reinforced specimens on both sides is lower than that of one side reinforcement. The strenght is higher with CFRP of high modulus and reinforced on both sides.	
6	[Colombi 2003]	Steel specimens hollowed at the centre with 2 cracks, representing the rivets holes, reinforced with 2 CFRP stripes post tensioned on both sides.	Tensile test	Steel specimens: 1000mmx300mmx10mm. SikaCarboDur stripes S512: 500mmx 50mmx1.2mm. Post tensioned stripes with high module: 500mmx50mmx1.4mm . Adhesive: 500x50x0.3mm.	Steel: E=210 Gpa. E(SikaCarboDur S512)=155 Gpa. E(SikaCarboDur M614)=210GPa. Adhesive E=714 Mpa	- not reinforced - reinf. on both side - reinf. on both side with post tensioned CFRP - reinf. on both side with post tensioned CFRP	80 80 80 80	300.000 900.000 1.800.000 6.000.000	A detachment on the interface adhesive-steel has been observed and the adhesive higher thickness has been demontrated to grow up the shear deformation on the adhesive layer.	
7	[Taljsten 2009]	5 metal specimens hollowed provided with two initial cracks	Tensile test	Steel specimens: 670x205mm. FRP: 400x50x1.4mm	Steel: Fy=294 Mpa, Ft=382 Mpa. CFRP E50C: Ft=2000Mpa, E=155 Gpa. CFRP M50C: Ft=2500Mpa, E=260 Gpa. Adhesive BPE 567: shear strenght=52.5 Mpa, E=6.24 Gpa, G=2.4 Gpa. Adhesive Pox SK 41: shear strenght==81 Mpa, E=9.87 Gpa, G=3.8 Gpa	A1: not reinforced B1: E-CFRP S adhesive B2: E-CFRP S adhesive B3: E-CFRP S adhesive B4: E-CFRP S adhesive C1: B CFRP M adhesive D1: CFRP E post tensioned adhesive D3: CFRP E post tensioned adhesive	97,50 97,50 97,50 97,50 97,50 97,50 97,50	470.000 1.150.000 1.470.000 1.510.000 1.250.000 1.760.000 3.780.000 4.930.000	The detachment lies on the interface steel-FRP.	

	Riferimento	Campioni	Tipo di test	Dim. materiali	Caratteristiche materiali	Dettaglio	Δσ [Mpa]	N	Note sulla rottura	Immagine esemplificativa
8	[Matta 2005]	Steel bars A36, single jointed (SDS) or double jointed (DDS), reinforced with FRP on both sides.	Tensile test	SDS: 610 mm long, 24.8 mm large, 7.9 mm thick. DDS: 410 mm long, 1.33 mm thick.	Steel: Fy=305.7 Mpa, Fu=445.2 MPa and E=196 Gpa. Carbodur ultimate strenght=2,990 MPa, longitudinal module=165.7 Gpa. Sikadur adhesive 30 tensile strenght=24.8 MPa, E=4.5 Gpa.	SDS DDS	83 73	1.000.000 1.000.000	Provini confrontati con categorie di dettagli saldati 36° e 50° EC3. SDS: distacco dell'interfaccia acciaio/adesivo. DDS: la rottura è una combinazione di un modo coesivo nell'adesivo e il distacco matrice/fibra dentro il composito.	
9	[Liu 2005]	12 plates jointed at the centre reinforced with CFRP on both side, normal and high module	Tensile test	Steel specimens: 50 mm large, 5 mm thick, 210 mm long. FRP: 80+60mm, 80+40mm	Provini c'acciaio: E=195 Gpa, ft = 484 Mpa. Mbrace CF 130: E = 240 Gpa, ft = 3800 Mpa. Mbrace CF 530 : E = 640 Gpa, ft = 2650 Mpa.	N40-1 N60-1 H40-1 H60-1	90 55 95 64	800.000 1.000.000 523.000 1.500.000	ID = failure to detachment in the interface. FB = fiber failure.	
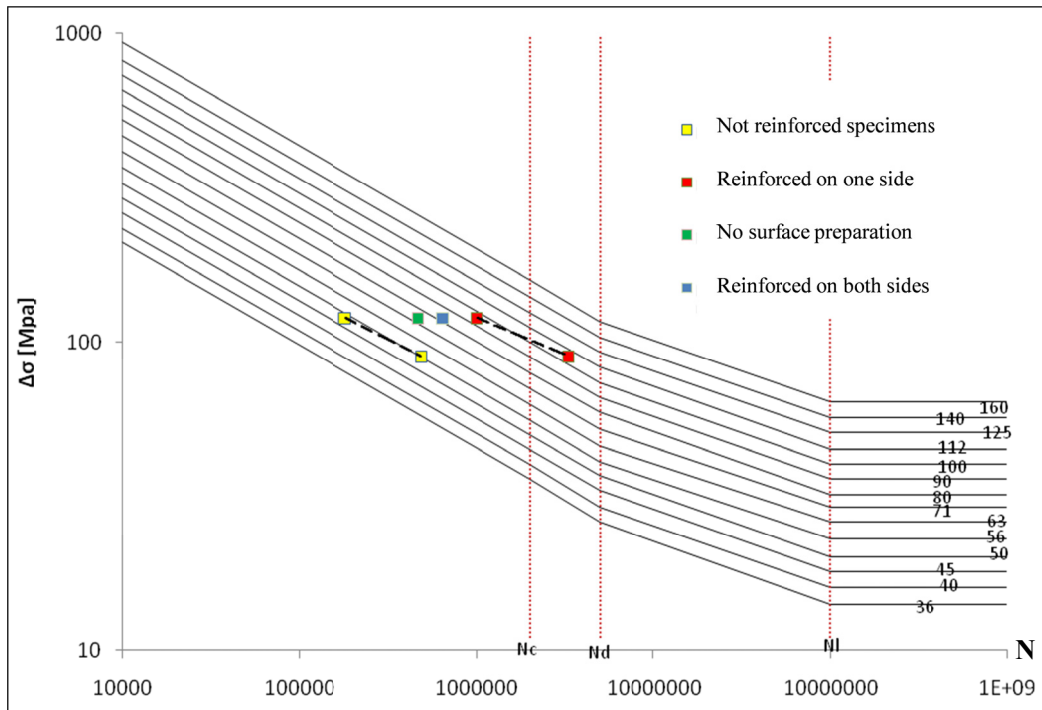
Figure 3. Fatigue failures and curve of FRP reinforced steel plates subjected to tension forces

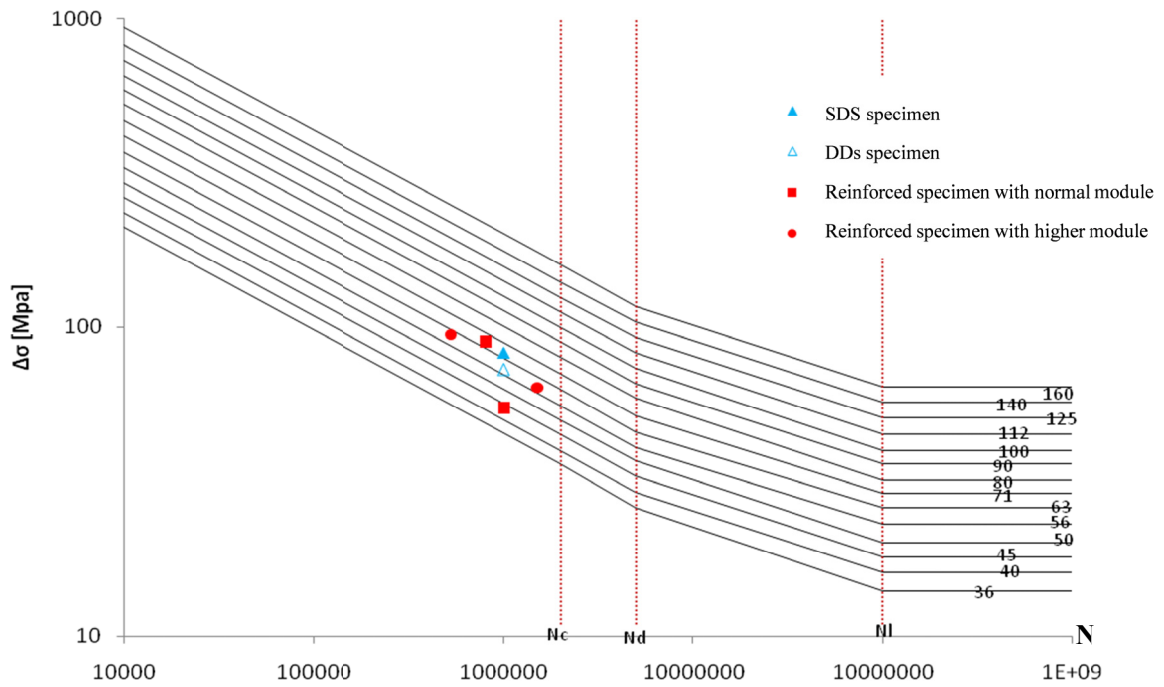Figure 4. Fatigue failures of double strap FRP reinforced steel plates subjected to tension forces
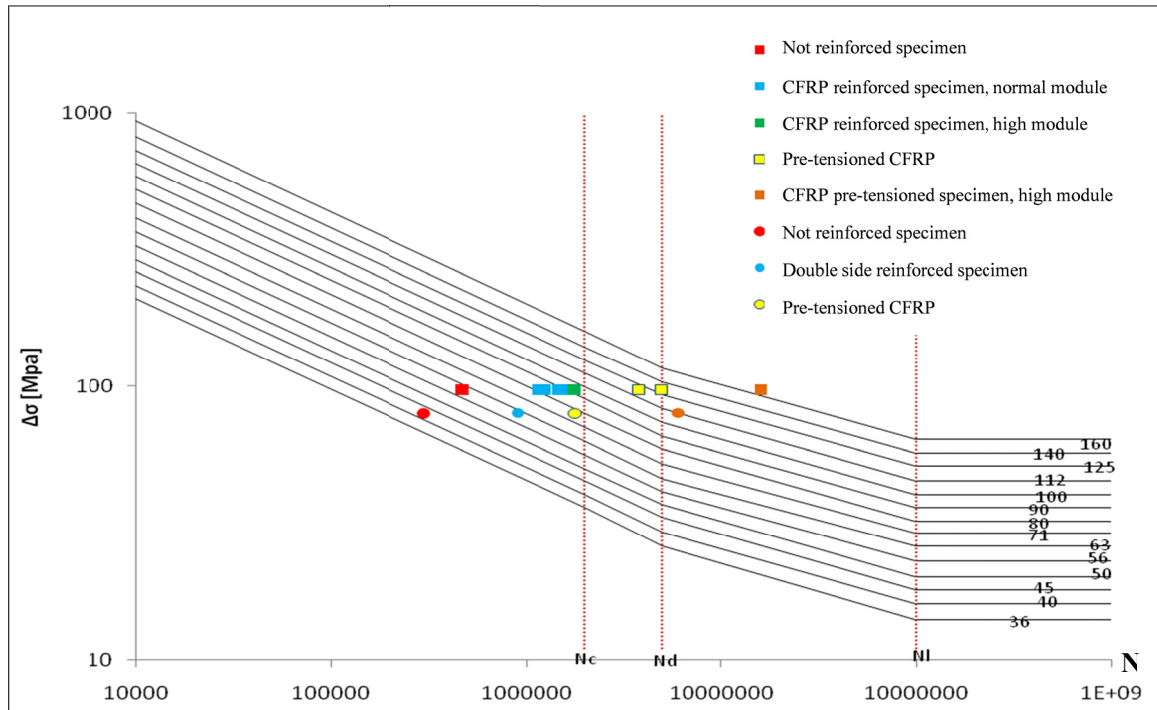
Figure 5. Fatigue failures of existing plates with FRP reinforcement subjected to tension forces
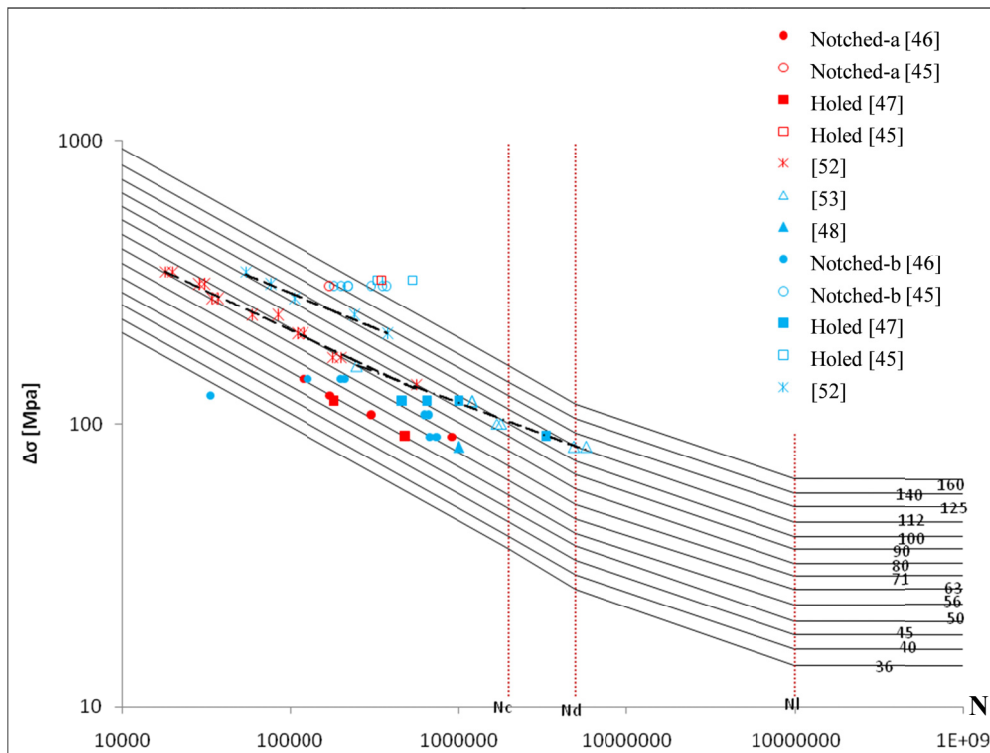

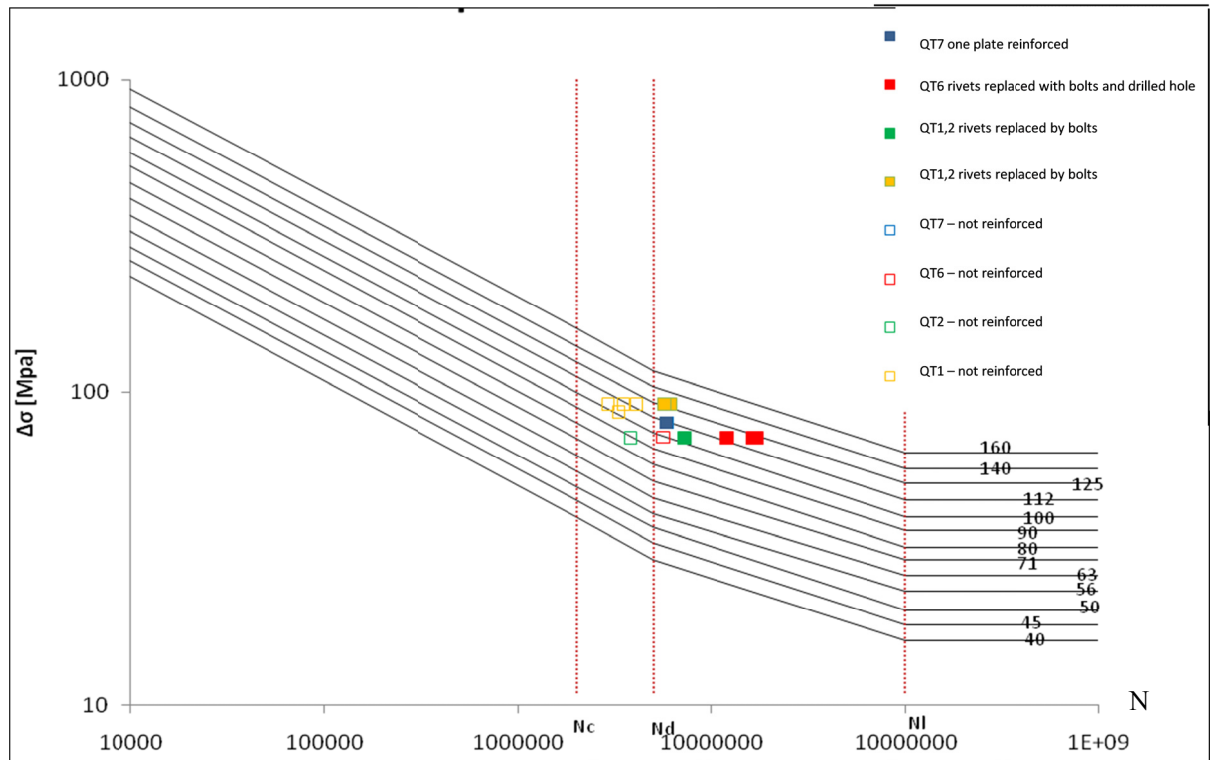Figure 6. Fatigue failures and curves of different tests

Figure 7. Fatigue failures of steel beam with traditional tecniques subjected to bending
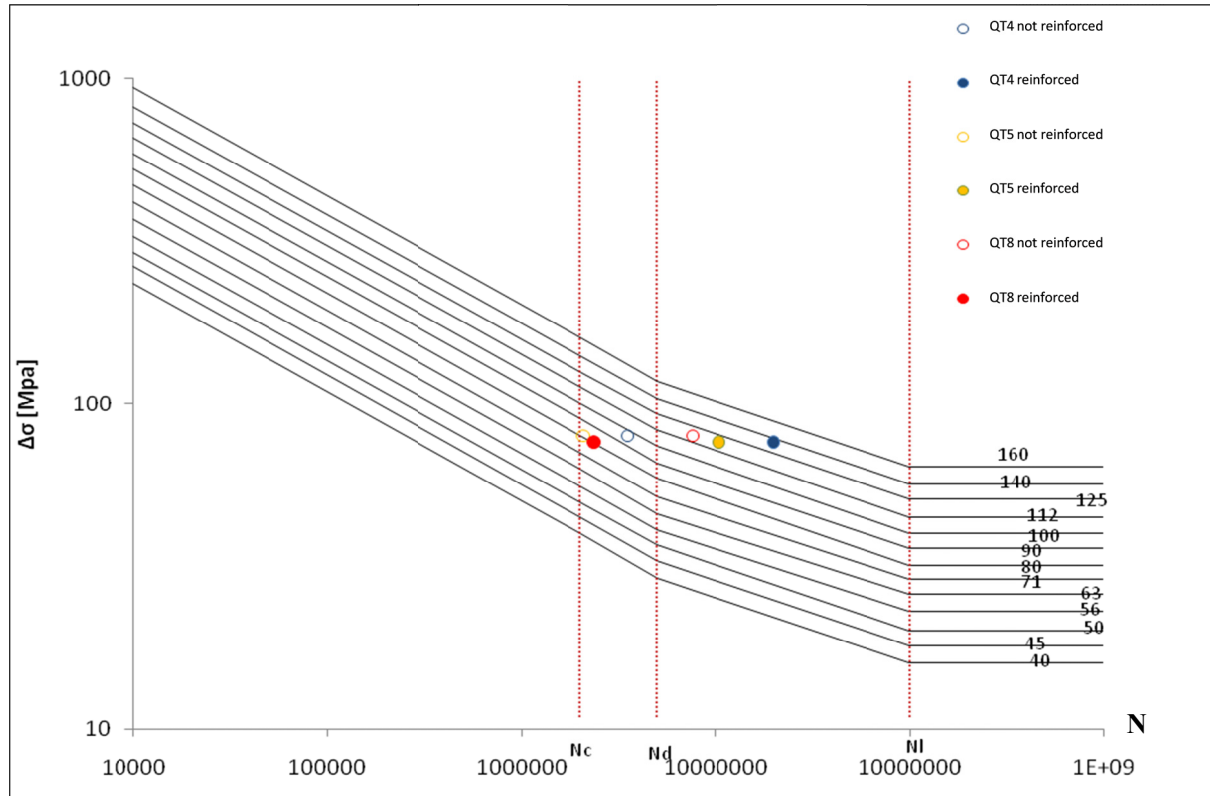

Figure 8. Fatigue failures of steel beam with CFRP reinforcing subjected to bending

5. Conclusions

This study deals with the analysis and comprehension of the fatigue behaviour of steel bridge joint, reinforced with FRP materials. The actual state of art of FRP strengthening is presented and discussed in the first part of this paper, while in the second part experimental tests coming from literature are presented and analyzed. It should be noted that, although a sufficient number of experimental tests have been found in literature, FRP reinforcing experimentations are more frequent for reinforced concrete than for steel structures, because up to now steel engineering have privileged steel to steel interventions rather than FRP rehabilitations. The use of FRP offers advantages compared to traditional techniques, because the FRP has a high strength / weight ratio, higher than that of steel; moreover they are very resistant against corrosion, if not entirely free from this, and extremely manageable. However, they are not sufficiently resistant to fire: FRP loses its strength and stiffness properties with temperature and the degradation in FRP properties is faster as compared to concrete or steel since the properties of the FRP matrix start to deteriorate even at modest temperature; also, bond degradation is another concern with externally bonded FRP. The analysis of experimental research found in the literature for the fatigue behavior of FRP reinforced steel detailing has proven the effectiveness of the composite material to extend the fatigue life of steel structures and decrease the growth of cracks. Two main types of fatigue testing has been analyzed, tensile and bending tests. Moreover a distinction has been made for test on individual steel plates reinforced with FRP, test on two samples of steel combined with FRP on both sides and the third is the tests on beams reinforced with FRP at the bottom. Failures data and their correlation were obtained from experimental data and compared to SN curves. Details reinforced with FRP belong to categories higher than those mentioned in EC3: in fact, for e.g., the graph of the fatigue curves of small-scale details of specimens reinforced virgin belong to the category 102, the notched specimens at 75, drilled to test 109, the old metal reinforced with 90 and those pre-tensioned to 125. From this analysis were then derived for the design of appropriate observations. The results obtained have shown a generally better fatigue behavior of steel reinforced detail than non-reinforced. Concerning the categories of Eurocode 3-1-9, reinforced details improved the ductility of the joint and demonstrated an higher life if compared to those of welded details, and it was also noted that the fatigue life of the steel specimens that have drilled holes is extended if compared to that of steel specimens which have cracks. This could be confirmed by an analysis of linear elastic fracture mechanics: the propagation of a crack can be described by a law of Paris as known, which takes into account the amplitudes of the effective stress intensity factors, which confirms that the speed of propagation of the crack will be higher and the fatigue life minor. The type of failure in almost all cases appears to be the detachment of the CFRP from the steel surface, without damaging the composite. This demonstrates the high capacity of resistance of fiber-reinforced composite and the weakest point of the system is the interface adhesive/steel. Finally it was found that the failure modes depend on the modulus of elasticity of the CFRP, the type of adhesive and the thickness of the adhesive.

Acknowledgment

References

ACI Committee 440. (2002). *Guide for the design and construction of externally bonded FRP systems for strengthening concrete structures (ACI 440.2R-02)*. American Concrete Institute, Farmington Hills, MI, USA.

Bassetti, A. (2001). Lamelles Precontraintes en Fibres Carbone pour le Renforcement de Ponts Rivet es Endomma ees par Fatigue (in French), Ph.D. Thesis no. 2440. Swiss Federal Institute of Technology, EPFL, Lausanne, 2001.

Bocciarelli, M., Colombi, P., Fava, G., & Poggi, C. (2009). Fatigue performance of tensile steel members strengthened with CFRP plates. *Composite Structures, 87*(2009), 334-343. http://dx.doi.org/10.1016/j.compstruct.2008.02.004

Buyukozturk, O., Gunes, O., & Karaca, E. (2004). Progress on understanding debonding problems in reinforced concrete and steel members strengthened using FRP composites. *Construction and Build Mater, 18*, 9-19. http://dx.doi.org/10.1016/S0950-0618(03)00094-1

Cadei, J. M. C., Stratford, T. J., Hollaway, L. C., & Duckett, W. G. (2004). *Strengthening metallic structures using externally bonded fibre-reinforced polymers*. Design Guide, London, ISBN 0-86017-595-2.

CNR Italian Advisory Committee on Technical Recommendations for Construction. (2004). *Guide for the design and construction of externally bonded FRP systems for strengthening existing structures*. Materials, RC and PC structures.

CNR Italian Research Council, Italian Advisory Committee on Technical Recommendations for Construction. (2005). *Guidelines for the design and construction of externally bonded FRP systems for strengthening existing structures. Preliminary study. Metallic structures (CNR-DT 202/2005).* Italian Research Council, Rome, ItalyMasonry structures (CNR-DT 200/2004). Italian Research Council, Rome, Italy.

Colombi, P. (2006). Reinforcement delamination of metallic beams strengthened by FRP strips: fracture mechanics based approach. *Eng. Fract. Mech., 73,* 1980-1996. http://dx.doi.org/10.1016/j.engfracmech.2006.03.011

Deng, J., Lee, M. M. K., & Moy, S. S. J. (2004). Stress analysis of steel beams reinforced with bonded CFRP plate. *Composite Struct, 65*(2), 205-215. http://dx.doi.org/10.1016/j.compstruct.2003.10.017

EN 1993-1-1. (2005). Design of steel structures, Part 1-1: General rules and rules for buildings. CEN, European committee for standardization, Bruxelles, Belgium.

Fib Task Group 9.3. (2001). Externally bonded FRP reinforcement for RC structures. fib bulletin 14, Lausanne, Switzerland.

Jao, Y., Teng, J. G., & Chen, J. F. (2005). Experimental study on FRP-to-concrete bonded joints. *Composites part B: Engineering, 36*(2), 99-113. http://dx.doi.org/10.1016/j.compositesb.2004.06.001

Jones, S. C., & Civjan, S. A. (2003). Application of fibre reinforced polymer overlays to extend steel fatigue life. *Journal of Composites for Construction, ASCE7*(4), 331-338. http://dx.doi.org/10.1061/(ASCE)1090-0268(2003)7:4(331)

Lenwari, A., Thepchatri, T., & Albrecht, P. (2005). Flexural response of steel beams strengthened with partial-length CFRP plates. *J. Compos Construct, 9*(4), 296-303. http://dx.doi.org/10.1061/(ASCE)1090-0268(2005)9:4(296)

Lenwari, A., Thepchatri, T., & Albrecht, P. (2006). Debonding strength of steel beams strengthened with CFRP strips. *J. Compos. Construct, 10*(1), 69-78. http://dx.doi.org/10.1061/(ASCE)1090-0268(2006)10:1(69)

Liu, H. B., Zhao, X. L., & Al-Mahaidi, R. (2005). The effect of fatigue loading on bond strength of CFRP bonded steel plate joints. International symposium on bond. Proceedings of the International Symposium on Bond Behaviour of FRP in Structures (BBFS 2005).

Lua, X. Z., Teng, J. G., Yea, L. P., & Jiang, J. J. (2005). Bond-slip models for FRP sheets/platesbonded to concrete. *Engineering Structures, 27*(6), 920-937. http://dx.doi.org/10.1016/j.engstruct.2005.01.014

Matta, F., & Karbhari, V. M. (2005). Tensile response of steel/CFRP adhesive bonds for the rehabilitation of civil structures. *Struct. Eng. Mech., 20*(5), 589-608.

Miller, T. C., Chajes, M. J., Mertz, D. R., & Hastings, J. N. (2001). Strengthening of a steel bridge girder using CFRP plates. *J. Bridge Eng., 6*(6), 514-522. http://dx.doi.org/10.1061/(ASCE)1084-0702(2001)6:6(514)

Monfared, A., Soudki, K., & Walbridge, S. (2008). CFRP reinforcing to extend the fatigue lives of steel structures. Fourth International Conference on FRP Composites in Civil Engineering (CICE2008) 22-24, July 2008, Zurich, Switzerland.

Moy, S. S. J. (2001). FRP composites: life extension and strengthening of metallic structures. ICE design and practice guide. Thomas Telford, London, ISBN 0-7277-3009-6 Pellegrino C, Modena C (2002) FRP shear strengthening of RC beams with transverse steel reinforcement. *J Compos Construct, 6*(2), 104-111.

Moy, S. S. J., & Nikoukar, F. (2002). Flexural behaviour of steel beams reinforced with carbon fibre reinforced polymer composite. In: Shenoi RA, Moy SSJ, Hollaway LC (eds) Advanced polymer composites for structural applications in construction. Proceedings of ACIC2002, inaugural international conference on the use of advanced composites in construction. Thomas Telford, London, pp. 195-202.

Moy, S. S. J., & Lillistone, D. (2006). Strengthening cast iron using FRP composites. *Proc Inst Civil Eng (ICE) Struct Build, 159*(6), 309-318. http://dx.doi.org/10.1680/stbu.2006.159.6.309

Moy, S. S. J., & Bloodworth, A. G. (2007). Strengthening a steel bridge with CFRP composites. *Proc Inst Civil Eng (ICE) Struct Build, 160*, 81-93. http://dx.doi.org/10.1680/stbu.2007.160.2.81

Pellegrino, C., Pipinato, A., & Modena, C. (2011). A simplified management procedure for bridge network maintenance. *Structure and Infrastructure Engineering, 7*(5), 341-351. http://dx.doi.org/10.1080/15732470802659084

Pipinato, A. (2010). Step level procedure for remaining fatigue life evaluation of one railway bridge. *Baltic Journal of Road and Bridge Engineering, 5*(1), 28-37. http://dx.doi.org/10.3846/bjrbe.2010.04

Pipinato, A. (2011). Assessment of existing bridges: Safety and security issues [Problemi di sicurezza nelle valutazioni strutturali di ponti esistenti]. *Ingegneria Ferroviaria, 66*(4), 355-371.

Pipinato, A. (2012c). Coupled safety assessment of cable stay bridges. *Modern Applied Science, 6*(7), 64-78. http://dx.doi.org/10.5539/mas.v6n7p64

Pipinato, A., & Modena, C. (2010). Structural analysis and fatigue reliability assessment of the Paderno bridge. *Practice Periodical on Structural Design and Construction, 15*(2), 109-124. http://dx.doi.org/10.1061/(ASCE)SC.1943-5576.0000037

Pipinato, A., Molinari, M., Pellegrino, C., Bursi, O. S., & Modena, C. (2011a). Fatigue tests on riveted steel elements taken from a railway bridge. *Structure and Infrastructure Engineering, 7*(12), 907-920. http://dx.doi.org/10.1080/15732470903099776

Pipinato, A., Pellegrino, C., Bursi, O. S., & Modena, C. (2009). High-cycle fatigue behavior of riveted connections for railway metal bridges. *Journal of Constructional Steel Research, 65*(12), 2167-2175. http://dx.doi.org/10.1061/(ASCE)SC.1943-5576.0000037

Pipinato, A., Pellegrino, C., & Modena, C. (2008). Scheduled maintenance actions and residual life evaluation of railway infrastructures for metallic structure bridges [Interventi di manutenzione programmata e valutazione della vita residua delle infrastrutture ferroviarie da ponte a struttura metallica]. *Ingegneria Ferroviaria, 63*(2), 125-134.

Pipinato, A., Pellegrino, C., & Modena, C. (2010). Structural analysis of historical metal bridges in Italy. *Advanced Materials Research, 133-134*, 525-530. http://dx.doi.org/10.4028/www.scientific.net/AMR.133-134.525

Pipinato, A., Pellegrino, C., & Modena, C. (2011b). Fatigue assessment of highway steel bridges in presence of seismic loading. *Engineering Structures, 33*(1), 202-209. http://dx.doi.org/10.1016/j.engstruct.2010.10.008

Pipinato, A., Pellegrino, C., & Modena, C. (2012a). Assessment procedure and rehabilitation criteria for the riveted railway Adige Bridge. *Structure and Infrastructure Engineering, 8*(8), 747-764. http://dx.doi.org/10.1080/15732479.2010.481674

Pipinato, A., Pellegrino, C., Fregno, G., & Modena, C. (2012b). Influence of fatigue on cable arrangement in cable-stayed bridges. *International Journal of Steel Structures, 12*(1), 107-123. http://dx.doi.org/10.1007/s13296-012-1010-5.

Pipinato, A., Pellegrino, C., & Modena, C. (2012d). Life assessment of riveted bridges: an analytical approach. ICE-Bridge Engineering, UK Institution of Bridge Engineers. In print.

Rizkalla, S., Dawood, M., & Schnerch, D. (2008). Development of a carbon fiber reinforced polymer system for strengthening steel structures. *Compos Part A Appl. Sci. Manufact, 39*(2), 388-397. http://dx.doi.org/10.1016/j.compositesa.2007.10.009

Shaat, A., Schnerch, D., Fam, A., & Rizkalla, S. (2004). Retrofit of steel structures using Fiber-Reinforced Polymers (FRP): state-of-the-art. In: Transportation research board (TRB) annual meeting. CD-ROM (04-4063).

Shaat, A., & Fam, A. (2006). Axial loading tests on short and long hollow structural steel columns retrofitted using carbon fibre reinforced polymers. *Can J. Civil Eng., 33*(4), 458-470. http://dx.doi.org/10.1139/l05-042

Schnerch, D., Dawood, M., Rizkalla, S., & Sumner, E. (2007). Proposed design guidelines for strengthening of steel bridges with FRP materials. *Construct Build Mater, 21*, 1001-1010. http://dx.doi.org/10.1016/j.conbuildmat.2006.03.003

Schnerch, D., & Rizkalla, S. (2008). Flexural strengthening of steel bridges with high modulus CFRP strips. *J. Bridge Eng., 13*(2), 192-201. http://dx.doi.org/10.1061/(ASCE)1084-0702(2008)13:2(192)

Sen, R., Liby, L., Spillett, K., & Mullins, G. (1995). Strengthening steel composite bridge members using CFRP laminates. In: Proceedings of the 2nd international RILEM symposium, FRPRCS-2, Aug. 23-25, Ghent, Belgium.

Smith, S. T., & Teng, J. G. (2001). Interfacial stresses in plated beams. *Eng. Struct., 23*, 857-871. http://dx.doi.org/10.1016/S0141-0296(00)00090-0

Stratford, T. J., & Cadei, J. M. C. (2006). Elastic analysis of adhesion stresses for the design of a strengthening plate bonded to a beam. *Construct Build Mater, 20*(1-2), 34-45. http://dx.doi.org/10.1016/j.conbuildmat.2005.06.041

Taljsten, B. (1997). Strengthening of beams by plate bonding. *J. Mater Civil Eng., 9*(4), 206-212. http://dx.doi.org/10.1061/(ASCE)0899-1561(1997)9:4(206)

Täljsten, B., Skodborg, C. H., & Wittrup J. S. (2001). Strengthening of old metallic structures in fatigue with prestressed and non-prestressed CFRP laminates. *Construction and Building Materials, 23*(2009), 1665-1677.

Tavakkolizadeh, M., & Saadatmanesh, H. (2003a). Strengthening of steel-concrete composite girders using carbon fiberreinforced polymer sheets. *J. Struct. Eng., 129*(1), 30-40. http://dx.doi.org/10.1061/(ASCE)0733-9445(2003)129:1(30)

Tavakkolizadeh, M., & Saadatmanesh, H. (2003b). Repair of damaged steel-concrete composite girders using carbon fiber-reinforced polymer sheets. *J. Compos. Construct, 7*(4), 311-322. http://dx.doi.org/10.1061/(ASCE)1090-0268(2003)7:4(311)

Tavakkolizadeh, M., & Saadatmanesh, H. (2003). Fatigue strength of steel girders strengthened with carbon fiber reinforced polymer patch. *Journal of Structural Engineering, ASCE129*(2), 186-196. http://dx.doi.org/10.1061/(ASCE)0733-9445(2003)129:2(186)

Yuana, H., Teng, J. G., Seracinoc, R., Wud, Z. S., & Yaoa, J. (2004). Full-range behavior of FRP-to-concrete bonded joints. *Engineering Structures, 26*(5), 553-565. http://dx.doi.org/10.1016/j.engstruct.2003.11.006

Zheng, Y., Ye, L. P., & Lu, X. Z. (2006). Experimental Study on Fatigue Behaviour of tensile plates strenghtened with CFRP plates. Third International Conference on FRP Composites in Civil Engineering (CICE 2006).

Zhao, X. L., & Zheng, L. (2007). State-of-the-art review on FRP strengthened steel structures. *Eng. Struct., 29*, 1808-1923. http://dx.doi.org/10.1016/j.engstruct.2006.10.006

Multi-scale Analysis of Void Closure for Heavy Ingot Hot Forging

Xiaoxun Zhang[1], Fang Ma[2], Kai Ma[1] & Xia Li[1]

[1] School of Materials Engineering, Shanghai University of Engineering Science, Shanghai, China

[2] College of Automotive Engineering, Shanghai University of Engineering Science, Shanghai, China

Correspondence: Xiaoxun Zhang, School of Materials Engineering, Shanghai University of Engineering Science, 333 Longteng Road, Songjiang District, Shanghai 201620, China. Tel: 86-139-1779-2182. E-mail: xx.zhang.cn@gmail.com

Abstract

A multi-scale model towards simulating the void closure during hot forging was introduced in the present study and the derived void evolution model was programmed into commercial code DEFORM to simulate the void closure behavior in heavy ingot during the upsetting and blocking processes. From the simulation results, it can be concluded that: (1) the cymbal-shaped die is good at closing voids not only near the die but also around the axis of the ingot when the reduction is 36% in the upsetting process, (2) the effectiveness of V-shaped die is the best for consolidating voids around the axis and near the die when the reduction is 24% in a single blocking process, however, at about 22% reduction after 90° rotate, the void-closed regions become quite large around the axis both for flat dies and FML dies and the loads are much lower than those of V-shaped dies. Upon that, void closure in multi-stroke and multi-pass forging was also discussed. The multi-scale model and simulation results provide valuable sources of reference for design and optimization of die shapes and pass schedules for heavy ingots during hot working processes.

Keywords: multi-scale, modeling and simulation, void closure, heavy ingot, hot forging

1. Introduction

Internal defects such as shrinkage cavities and porosity are usually generated in heavy ingots during solidification in steel casting. These tiny voids must be closed up in the subsequent hot forging process to ensure high quality of the product (Dudra & Im, 1990). Because of its importance, void closure has been studied for more than 30 years. Various methods, such as physical simulation and experimental study (Chaaban & Alexander, 1976), upper bound analysis (Ståhlberg, 1986), finite element (FE) method (Tanaka et al., 1986; Park & Yang, 1997; Jiang et al., 2005; Lee et al., 2011) are used to develop predictive measures for void closure in heavy ingot during hot forging. However, since a large number of small voids usually exist in heavy ingot and the volume of the void is extremely small compared with the heavy ingot, modeling void closure in heavy ingot during hot forging is a multi-scale problem, and it is very difficult to find predictive measures based only on macro-mechanics.

In this study, a meso-mechanics approach is employed to deal with the multi-scale problem. A cell model is adopted to study the deformation rate of the void and a theoretical model towards simulating the void closure during hot forging is introduced. The derived void evolution equation is programmed into commercial software DEFORM to simulate the behavior of void closure in heavy ingot during the upsetting and blocking processes. The relative void volume, which is defined as the ratio of current void volume to initial void volume, is calculated during the deformation of heavy ingots in the hot working processes. Relative void volume and its distribution are used to evaluate the effectiveness of different die shapes and the processes for consolidating internal voids. Both upsetting and blocking processes for heavy ingots are studied. Upon that, void closure in multi-stroke and multi-pass forging is also discussed.

2. A Multi-scale Model for Void Closure

2.1 Cell Model and Void Evolution

A cell model is adopted to deal with the multi-scale problem of void closure and to analyze the evolution of void. The cell model which includes matrix and void is shown in Figure 1 with volume V_c and outer surface S_c, and

subjected to remote (macroscopic) uniform stress $\boldsymbol{\Sigma}$. The matrix material is assumed to be isotropic and incompressible. The constitutive relation of the matrix during hot forging can be given as (Cocks, 1989; Duva & Hutchinson, 1984)

$$\dot{\boldsymbol{\varepsilon}} = \frac{3}{2} \frac{\dot{\varepsilon}_0}{\sigma_0} \left(\frac{\sigma_e}{\sigma_0} \right)^{n-1} \boldsymbol{\sigma}' \tag{1}$$

where $\dot{\varepsilon}_0$ and σ_0 are reference strain-rate and stress respectively, n ($1 \le n < +\infty$) is the Norton exponent, $\boldsymbol{\sigma}'$ is the stress deviator and $\sigma_e = \left(\frac{3}{2} \boldsymbol{\sigma}' : \boldsymbol{\sigma}' \right)^{1/2}$ is the effective stress. An initially spherical void is contained in the cell.

The volume of the void is V and the void surface is S. The void volume fraction is assumed small and void interaction is ignored. The aim is to obtain the velocity field in the matrix of the cell, and then the evolution of the void can be determined by this velocity field.
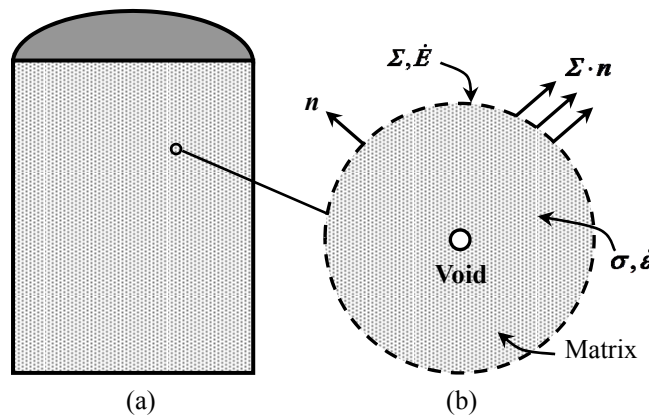
Figure 1. (a) Heavy ingot: comprised of a matrix and a dilute dispersion of traction-free voids and (b) Cell model: volume V_c, outer surface S_c, uniform tractions $\boldsymbol{\Sigma} \cdot \boldsymbol{n}$ over S_c

To obtain an evolution equation of the void, a Rayleigh-Ritz procedure based on Hill's minimum principle (Hill, 1956) for the velocities is developed. The trial local velocity and strain-rate within the matrix are expressed as

$$\boldsymbol{v} = \boldsymbol{v}^0 + \tilde{\boldsymbol{v}}, \quad \dot{\boldsymbol{\varepsilon}} = \dot{\boldsymbol{E}}^0 + \dot{\tilde{\boldsymbol{\varepsilon}}} \tag{2}$$

where $\boldsymbol{v}^0$ and $\dot{\boldsymbol{E}}^0$ are the uniform velocity and strain-rate due to $\boldsymbol{\Sigma}$ in the absence of the void. Then, among all additional velocity fields $\tilde{\boldsymbol{v}}$, the actual velocity field minimizes the functional (Duva & Hutchinson, 1984)

$$F(\tilde{\boldsymbol{v}}) = \int_{V_m} [w(\dot{\boldsymbol{\varepsilon}}) - w(\dot{\boldsymbol{E}}^0) - \boldsymbol{\Sigma} : \dot{\tilde{\boldsymbol{\varepsilon}}}] dV - \int_S \tilde{\boldsymbol{v}} \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{n} dS \tag{3}$$

where V_m is the volume of the matrix material in the cell, S is the surface of the void, $\boldsymbol{n}$ is the unit normal to the surface of the void pointing into the matrix, and

$$w(\dot{\boldsymbol{\varepsilon}}) = \frac{n}{n+1} \dot{\varepsilon}_0 \sigma_0 \left(\frac{\dot{\varepsilon}_e}{\dot{\varepsilon}_0} \right)^{(n+1)/n} \tag{4}$$

in which $\dot{\varepsilon}_e = (\frac{2}{3} \dot{\boldsymbol{\varepsilon}} : \dot{\boldsymbol{\varepsilon}})^{1/2}$ is the local effective strain-rate. Note that $\lim_{\rho \to 0} \dot{\boldsymbol{E}} = \dot{\boldsymbol{E}}^0$, where $\rho = V / V_c$ is the void volume fraction and V is the volume of the void. Since the cell model considered for heavy ingots is infinite compared with the void, $\dot{\boldsymbol{E}} = \dot{\boldsymbol{E}}^0$ will be used throughout and

$$\dot{\boldsymbol{E}} = \dot{\boldsymbol{E}}^0 = \frac{3}{2}\frac{\dot{\varepsilon}_0}{\sigma_0}\left(\frac{\Sigma_e}{\sigma_0}\right)^{n-1}\boldsymbol{\Sigma}' \tag{5}$$

As the matrix material is incompressible, the stream function can be introduced to characterize the additional velocity field $\tilde{\boldsymbol{v}}$ and the unknown variables in the stream function can be determined by minimizing $F(\tilde{\boldsymbol{v}})$ in Eq.(3) (Budiansky et al., 1982; Lee & Mear, 1994). Once the velocity field $\boldsymbol{v}$ has been computed from the Rayleigh-Ritz procedure, the change-rate of void volume is given by

$$\frac{\dot{V}}{V} = \frac{1}{V}\int_S \boldsymbol{v}\cdot\boldsymbol{n}\,dS \tag{6}$$

2.2 The Multi-scale Model for Void Closure

Based on the calculations of the Ritz procedure mentioned above, the numerical solutions were obtained in a previous study by Zhang et al. (2009). These numerical results can be used to formulate an explicit expression of the multi-scale model for void closure, which is suggested as

$$\begin{cases} \dfrac{V}{V_0} = \exp\left\{\text{sign}(\Sigma_m)E_e\left[\dfrac{3}{2}\left(\dfrac{3}{2n}\left|\dfrac{\Sigma_m}{\Sigma_e}\right|+\dfrac{(n-1)(5n+2)}{5n^2}\right)^n + q_1\left|\dfrac{\Sigma_m}{\Sigma_e}\right|+q_2 E_e^2+q_3 E_e^4+q_4\right]\right\} & \text{for } \Sigma_e \neq 0 \\[4mm] \dfrac{V}{V_0} = \exp\left[\text{sign}(\Sigma_m)\dfrac{3\dot{\varepsilon}_0 t}{2}\left(\dfrac{3}{2n\sigma_0}|\Sigma_m|\right)^n\right] & \text{for } \Sigma_e = 0 \end{cases} \tag{7}$$

where $\Sigma_m \equiv \frac{1}{3}\text{tr}(\boldsymbol{\Sigma})$ and $\Sigma_e \equiv \left(\frac{3}{2}\boldsymbol{\Sigma}':\boldsymbol{\Sigma}'\right)^{1/2}$ are the remote mean stress and the remote effective stress, respectively, and $\dot{E}_e = \left(\frac{2}{3}\dot{\boldsymbol{E}}:\dot{\boldsymbol{E}}\right)^{1/2}$ is the remote effective strain-rate, Σ_m/Σ_e is the measure of stress triaxiality, q_1, q_2, q_3 and q_4 are four parameters and E_e is the macroscopic effective strain. In the present study, FE computations based on a cubic cell model are carried out to determine the values of q_1, q_2, q_3 and q_4 and the values of q_1, q_2, q_3 and q_4 are suggested and given in Table 1.

Table 1. The values of q_1, q_2, q_3 and q_4 in Eq. (7)

n	1	2	3	5	10	100
q_1	0.5048	0.4911	0.6016	1.1481	2.9132	6.5456
q_2	6.4675	0.8002	-0.6981	-4.2026	-11.6464	-15.3775
q_3	14.2610	53.8018	72.6397	108.2114	185.5622	324.4417
q_4	-0.3379	-0.2314	-0.1243	-0.2480	-0.6511	-1.9575

3. Multi-scale Simulation of Void Closure

Upsetting and blocking are two typical operations in the process of manufacturing heavy forgings. Upsetting is a process which increases the cross-section of the billet by compressing its length, while blocking is a process which reduces the cross-section of the billet and increases its length by repetitive side pressing and alternate ingot rotations. Both upsetting and blocking processes are often used to eliminate internal voids in heavy ingots. New die geometries which lower the press loads have also been investigated since the press load capacity of the forging machine is a limitation when forging heavy ingots.

In this section, the derived void evolution equation is programmed into commercial software DEFORM to simulate the behavior of void closure in heavy ingot during the upsetting and blocking processes. The relative void volume R_V, which is defined as the ratio of current void volume V to initial void volume V_0, is calculated during the deformation of heavy ingots in the hot working processes. Relative void volume and its distribution are used to evaluate the effectiveness of different die shapes and the processes for consolidating

internal voids.

3.1 Void Closures in Upsetting

Upsetting of a cylindrical ingot with concave sphere die, flat die, M-shaped die, convex sphere die and cymbal-shaped die are simulated in this subsection to reveal the effect of die shapes on void closure. The FE models for upsetting with different die types are shown in Figure 2 and the diameter of the cylindrical ingot is D =1000 mm and the height is H = 2000 mm. The results of R_V by applying the multi-scale method (MSM) in simulation are shown in Figure 3 at 36% reduction. Note that $R_V \rightarrow 0$ means void closure, whereas $R_V \rightarrow 1$ indicates void has not closed. From Figure 3, it is clear that: (1) There is a void-unclosed region near the die during upsetting with concave sphere die and flat die (Figure 3(a) (b)). The void-unclosed region produced by concave sphere die is the biggest and R_V approaches 1 in that region. The void-unclosed region produced by flat die is smaller than that by concave sphere die, but it is still a disadvantage to obtain high quality forgings. It indicates that, from a point of view of eliminating voids in heavy ingots, concave sphere die and flat die are not suitable for upsetting. (2) The void-unclosed regions produced by convex surface die are much smaller than those by concave surface die and flat die (Figure 3(c) (d) (e)). Comparing M-shaped die, convex sphere die and cymbal-shaped die with each other, it can be seen that there is a relatively big void-unclosed region produced by M-shaped die near the intersection of the convex and concave surface, and R_V around the axis is also not uniform. Convex sphere die and cymbal-shaped die, however, can consolidate the void near the die during upsetting. Moreover, cymbal-shaped die upsetting provides an ideal result for void closure not only near the die but also around the axis of the ingot (Figure 3(e)). (3) Since the void-unclosed region always exists near the die during upsetting with concave sphere die, flat die and M-shaped die (Figure 3(a) (b) (c)), it is very difficult to achieve the goal of eliminating void in heavy ingots even by multiple upsetting along the axis. However, the voids near the die and those around the axis of the ingot can be eliminated with convex sphere die and cymbal-shaped die (Figure 3(d) (e)), so that the goal of eliminating voids can be achieved by multiple upsetting along the axis with convex sphere die and cymbal-shaped die. (4) Since the metallurgy defects in heavy ingots scatter mainly along the axis, the cavity defects would be eliminated completely only when all of the voids around the axis are closed. Since cymbal-shaped die is good at closing voids not only near the die but also around the axis of the ingot, and the loads for cymbal-shaped die are also relatively small, the cymbal-shaped die is suggested to use in upsetting.

(a)   (b)   (c)

(d)   (e)
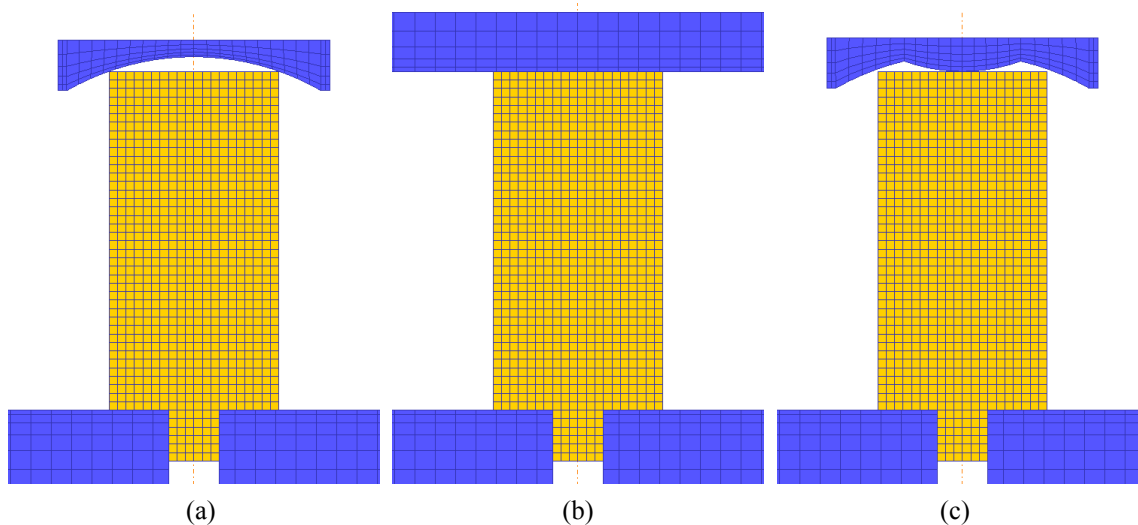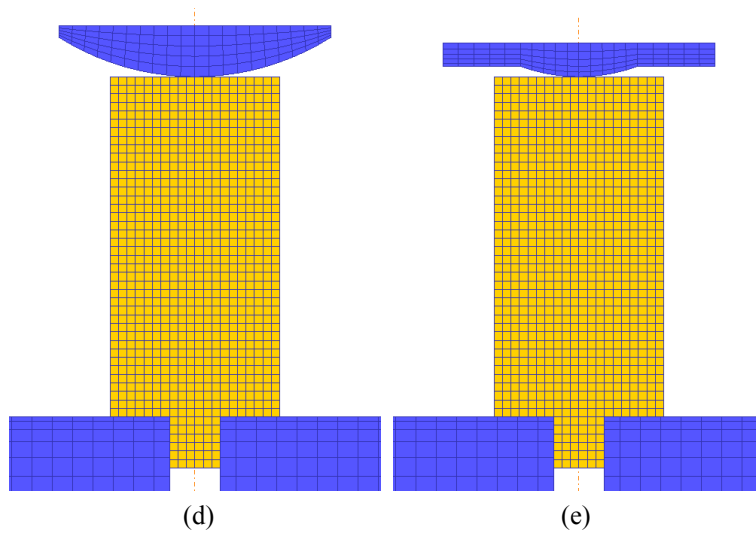
Figure 2. Finite element models for upsetting with different die types: (a) concave sphere die, (b) flat die, (c) M-shaped die, (d) convex sphere die and (e) cymbal-shaped die

(a)   (b)   (c)

(d)   (e)

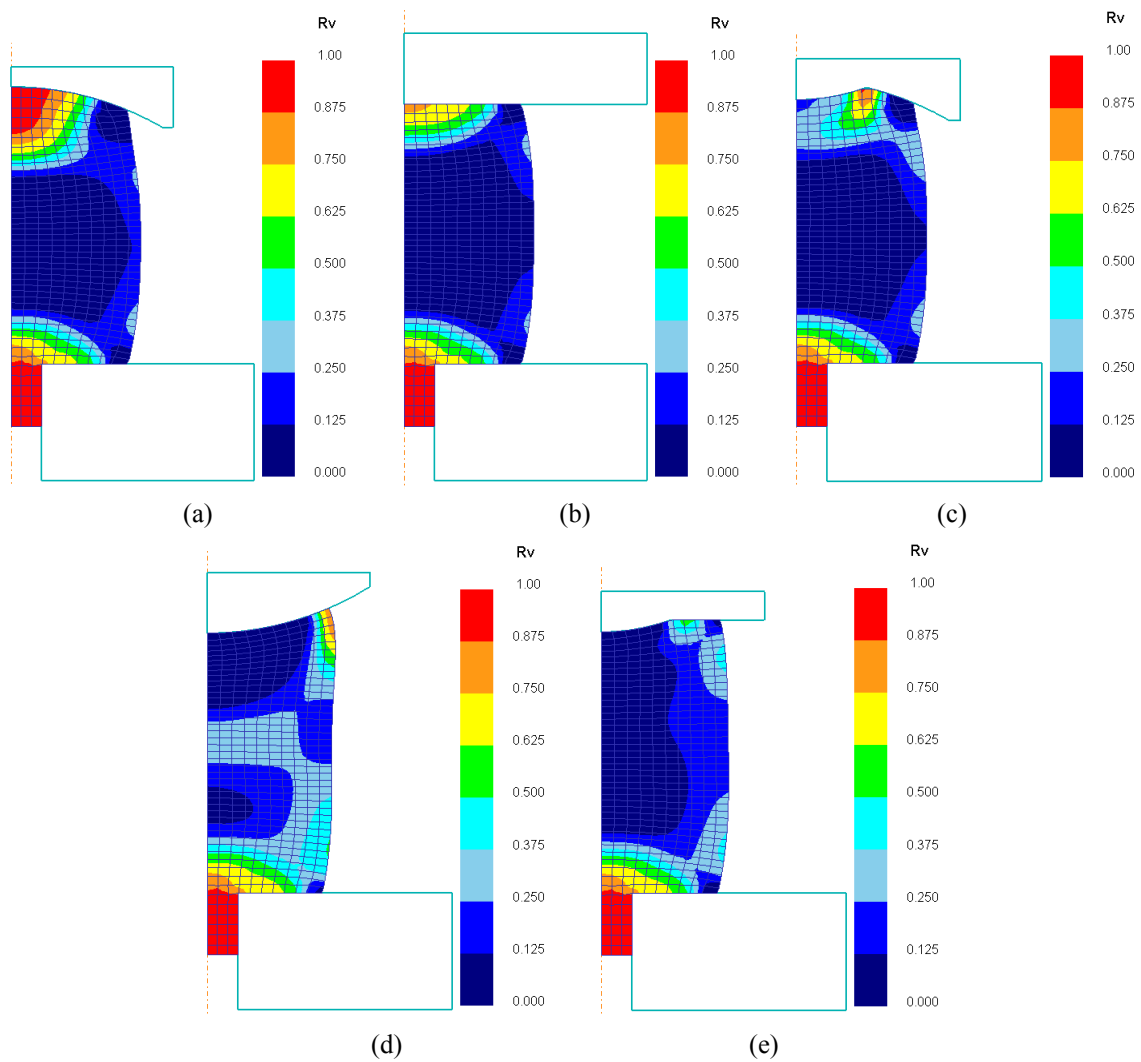Figure 3. Relative void volume for upsetting with different dies at 36% reduction: (a) concave sphere die, (b) flat die, (c) M-shaped die, (d) convex sphere die, (e) cymbal-shaped die

3.2 Void Closures in Blocking

Blocking of a cylindrical ingot with V-shaped die, flat die and FML dies are simulated in this subsection to evaluate the effectiveness of these dies for consolidating internal porosity. The FE models for blocking with different dies are shown in Figure 4 and the diameter of the cylindrical ingot is 1400 mm. The results of R_V at 24% reduction are shown in Figure 5, Figure 6(a) and Figure 7(a). Press loads required for different dies are shown in Figure 8. From Figures 4~8, it is found that: (1) The distribution of void-closed region is significantly influenced by die shapes. Void-closed region produced by V-shaped die is the largest and effectiveness of V-shaped die is the best for consolidating voids around the axis and near the die (Figure 5). Void-closed region produced by flat die is mainly distributed around the axis (Figure 6(a)) and void-closed region produced by FML dies is mainly centered on a small area around the axis (Figure 7(a)). (2) The loads for V-shaped die increase rapidly from the beginning of loading and keep a high growth rate, whereas the loads for flat die and FML dies are relatively much smaller (see Figure 8). The loads for flat die and FML dies are almost the same when the reduction is less than 13% (the reduction time is 6 s). The loads for flat die are higher than those for FML dies when the reduction is greater than 13%. It indicates that, from a point of view of eliminating voids in heavy ingots, the effectiveness of V-shaped die is the best but the loads required are also the highest. The effectiveness of flat die takes second place with lower loads. The FML dies are less effective in each stroke than other dies for consolidating void in heavy ingot, but the loads required for FML dies are also the smallest. If the press load of the forging machine is enough, then the pair of V-shaped dies would be a good choice for blocking. These results are in good agreement with the experimental and theoretical results in literatures. (3) One of the characteristics of blocking is that the billet can be pressed repetitively with alternate ingot rotations. The distributions of R_V at 24% reduction and those at 22% reduction after 90° rotate are displayed in Figure 6 and Figure 7 with flat dies and FML dies, respectively. At the first 24% reduction, the void-closed regions produced by flat die and FML dies are very small (see Figure 6(a) and Figure 7(a)). At about 22% reduction after 90° rotate, however, the void-closed regions become quite large around the axis both for flat dies and FML dies (see Figure 6(b) and Figure 7(b)), and the loads are much lower than those of V-shaped dies (Figure 8). The void-closed region near the die during blocking with FML dies is larger than that with flat dies. If press load of the forging machine is a constraint, then the FML dies would be a better choice. It also implies that the voids in heavy ingots would be closed by multi-stroke and multi-pass blocking with flat dies and FML dies.

(a)   (b)   (c)
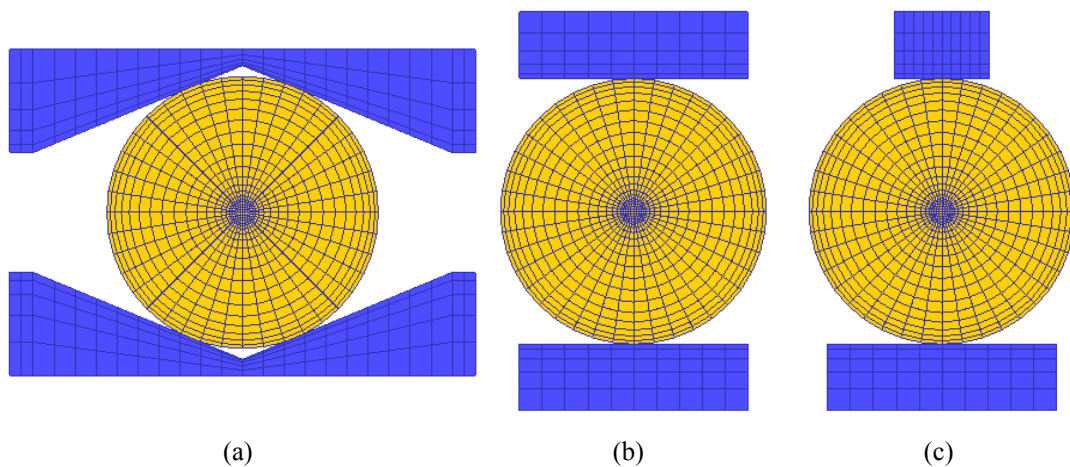
Figure 4. Finite element models for blocking with different die types: (a) 135°V dies, (b) flat dies, (c) FML dies
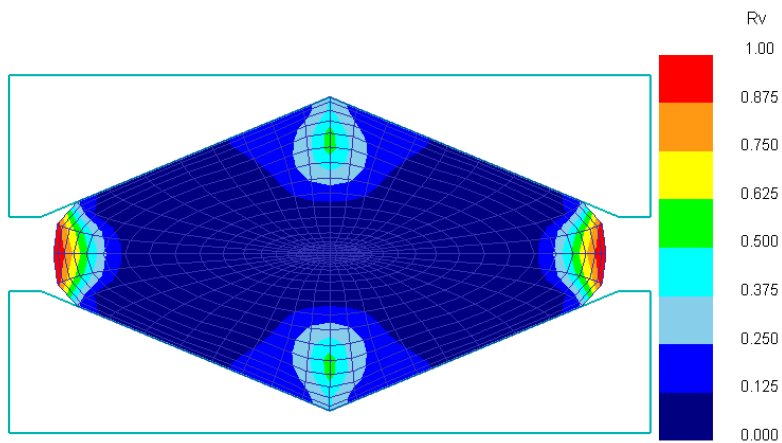
Figure 5. Relative void volume for 135° V dies blocking at 24% reduction

(a)   (b)
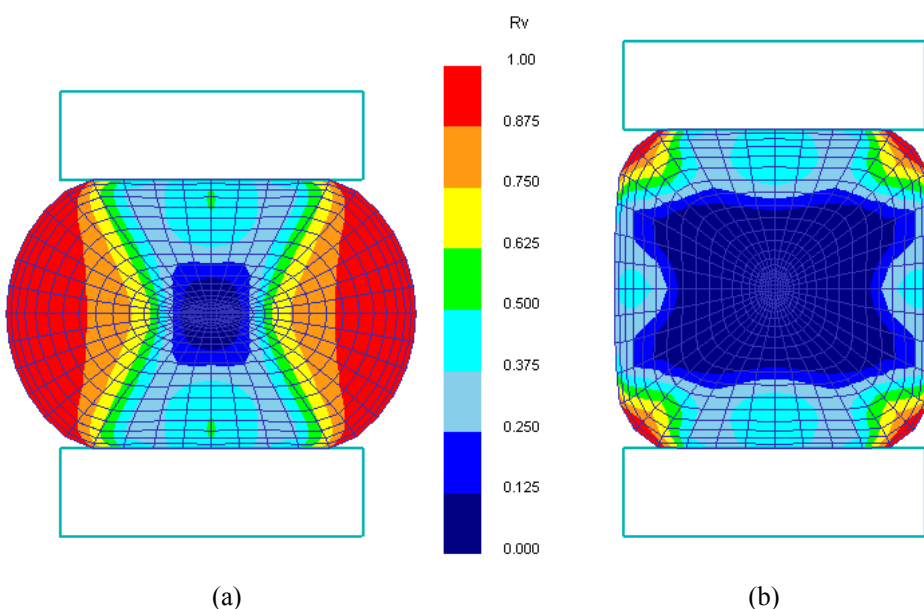
Figure 6. Relative void volume for flat dies blocking: (a) 24% reduction, (b) 22.2% reduction after 90° rotates

(a)   (b)
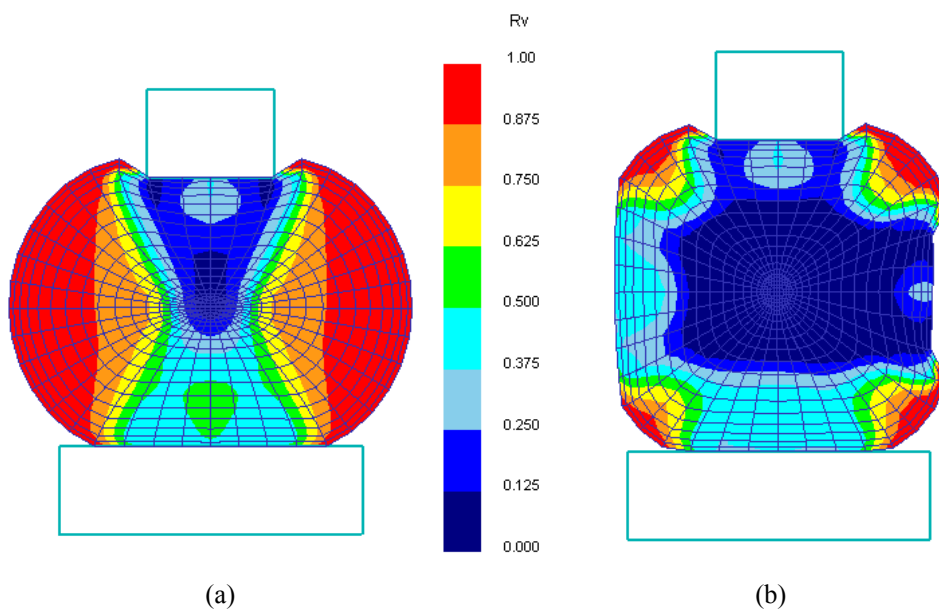
Figure 7. Relative void volume for FML dies blocking: (a) 24% reduction, (b) 22.6% reduction after 90° rotates
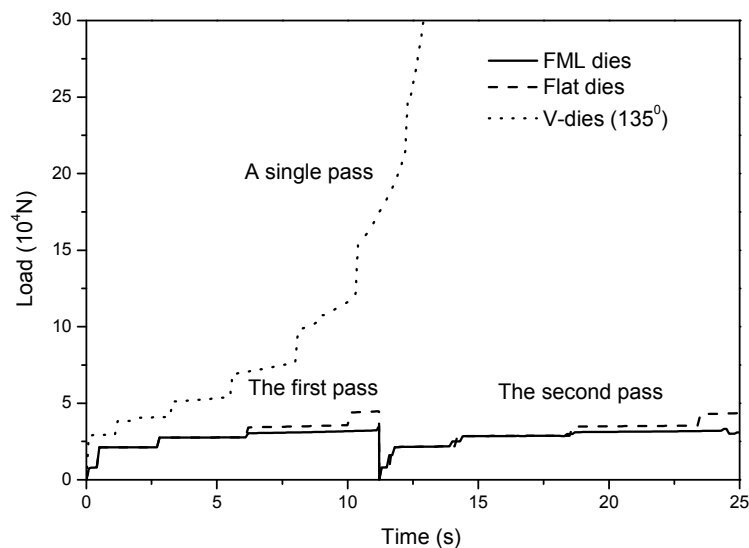
Figure 8. Loads for blocking with different dies

3.3 Multi-stroke and Multi-pass Forging

The forging process of heavy ingots usually needs a lot of strokes and many passes, therefore study on void closure in multi-stroke and multi-pass is very important. Only in this way, the course of forging can be evaluated in an all-round way and the process can be optimized as a whole.

Void closure in a square billet during multi-stroke blocking with a pair of flat dies is investigated in this subsection. The FE model of the square billet is shown in Figure 9 and its width is W =1000 mm, height is H = 1200 mm and length is L =3000 mm. The results of R_V after four strokes in the first pass are shown in Figure 10(a) and the reduction for each stroke is 25%. Figure 10(b) is a sliced plane view. After the first pass, the ingot is rotated 90° and five strokes are performed in the second pass. Figure 10(c) displays the results of R_V in the second pass. From Figure 9 and Figure 10, it can be seen that: (1) Looking from side, there are "X" shaped areas whose internal cavities are closed fairly well after four strokes in the first pass (see Figure 10(a)(b)). (2) The sliced plane view indicates that the voids in some local regions of the ingot has been closed after four strokes, nevertheless, continuous void-closed regions along the centerline of the ingot still have not appeared after one pass forging. (3) Since void-unclosed region always exists near the flat dies during every stroke, and void-unclosed region appears between strokes, the multi-pass blockings are needed to eliminate the internal voids in the ingot. Void-closed regions will be continuously produced by alternate press and rotate in multi-pass blocking (see Figure 10(c)). The goal of eliminating void would be achieved when these local void-closed regions are united as a whole in the ingot. (4) Application of the criterion for void closure in the CAE analysis makes it very convenient to evaluate and optimize various traditional forging processes, and provides a novel way for new process design in terms of elimination of voids in heavy ingots.
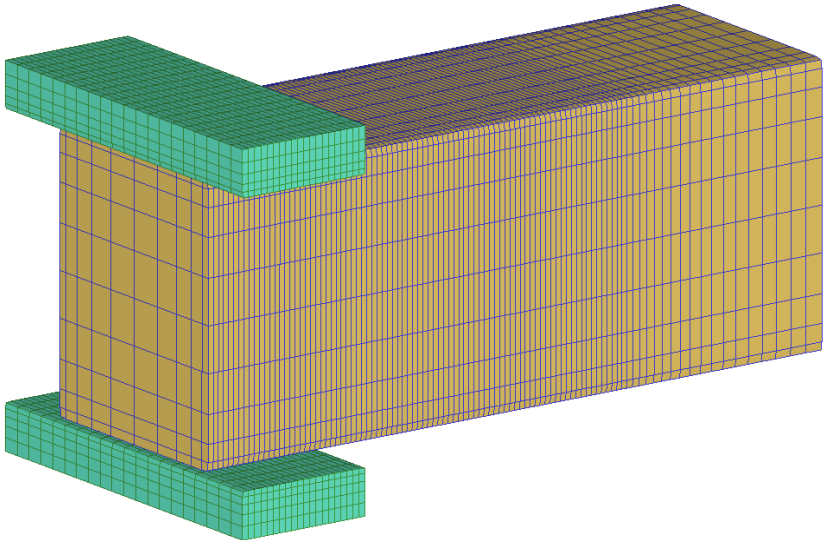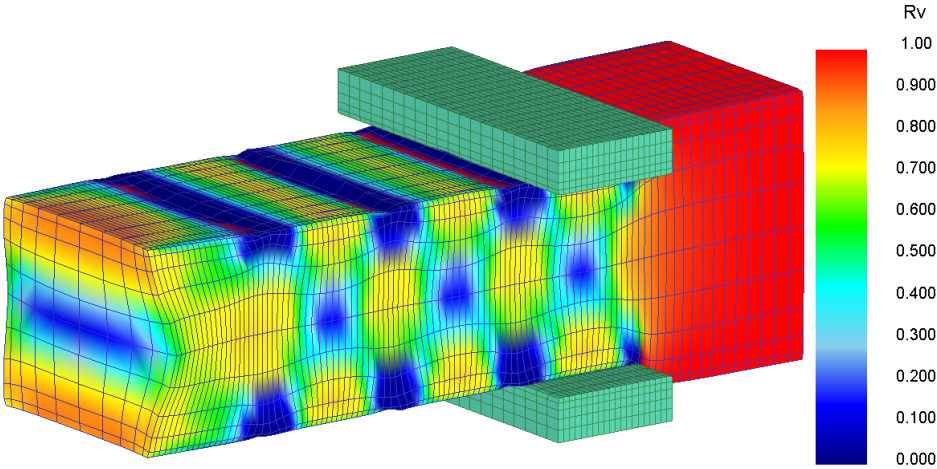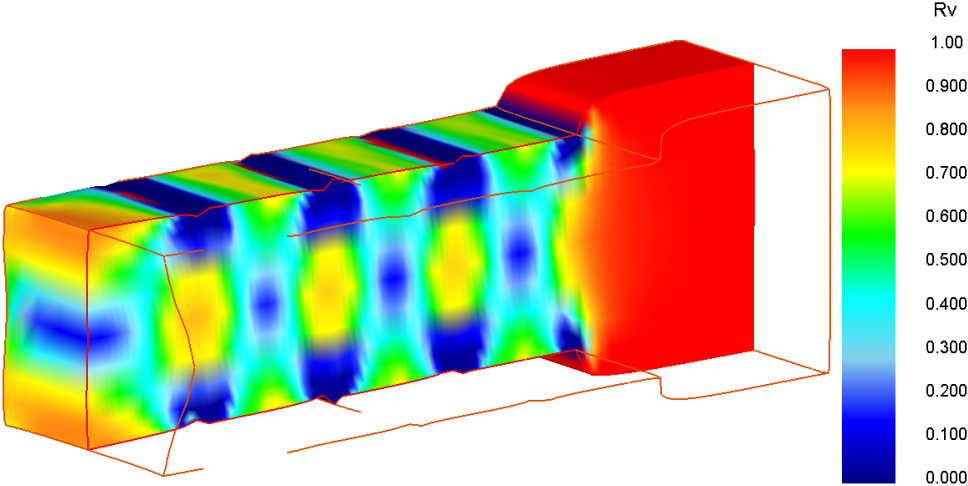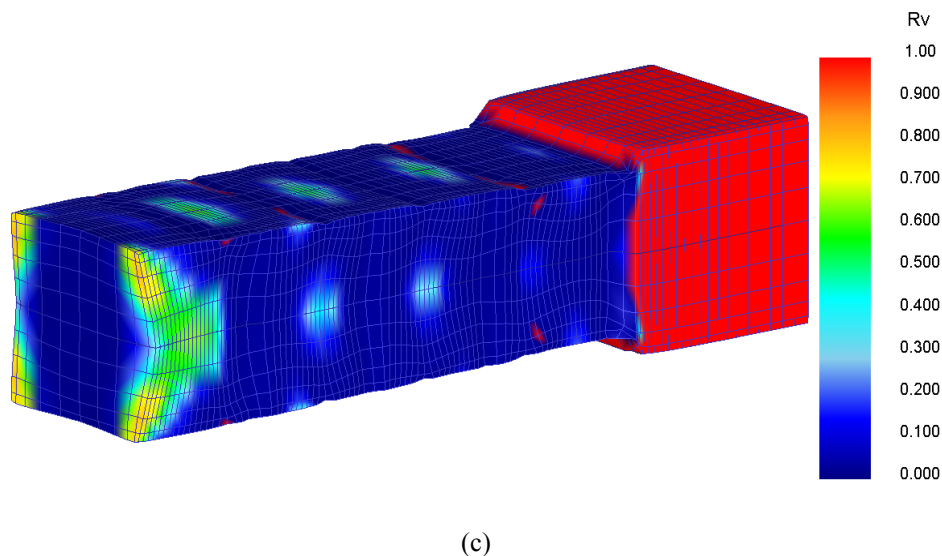
Figure 9. Finite element models for multi-stroke and multi-pass forging

(a)

(b)

(c)

Figure 10. Void closure in multi-stroke and multi-pass forging: (a) after four strokes in the first pass, (b) section plane view after four strokes in the first pass, and (c) after five strokes in the second pass

4. Conclusions

A multi-scale model towards simulating the void closure during hot forging is introduced in this study. A cell model, which includes a void-free matrix and a void, is adopted, and Rayleigh-Ritz procedure is used to study the deformation rate of the void. The derived void evolution model is programmed into commercial code DEFORM to simulate the void closure behavior in heavy ingot during the upsetting and blocking processes. It can be concluded that: (1) From a point of view of eliminating voids in heavy ingots, concave sphere die and flat die are not suitable for upsetting since a large void-unclosed region exists near the die. Since the cymbal-shaped die is good at closing voids not only near the die but also around the axis of the ingot, and the loads for cymbal-shaped die are also relatively small, the cymbal-shaped die is suggested to use in upsetting. (2) From a point of view of eliminating voids in heavy ingots, the effectiveness of V-shaped die is the best for blocking but the load required is also the highest. The effectiveness of flat die takes second place with lower loads. The FML dies are less effective in each stroke than other dies, but the loads required for FML dies are also the smallest. If the press load of the forging machine is enough, then the pair of V-shaped dies would be a good choice for blocking. If press load of the machine is a constraint, then the FML dies would be a better choice. (3) Since void-unclosed region exists near the flat die during every stroke, and void-unclosed region appears between strokes, the multi-pass blockings are needed to eliminate the internal voids in heavy ingots. Void-closed regions will be continuously produced by alternate press and rotate in multi-pass blocking. When the local void-closed regions are united as a whole in the ingot, the goal of eliminating void will be achieved. (4) By applying the multi-scale model for void closure in the CAE analysis, the optimal forging process, in terms of elimination of voids in heavy ingot, can be carried out and the schedule for multi-stroke and multi-pass can be arranged economically, and then high product quality would be obtained.

Acknowledgements

References

Budiansky, B., Hutchinson, J. W., & Slutsky, S. (1982). Void growth and collapse in viscous solids. In: Hopkins, H.G., Sewell, M.J. (Eds.), *Mechanics of Solids* (pp. 13-45), Pergamon Press, Oxford.

Chaaban, M. A., & Alexander, J. M. (1976). A Study of the Closure of Cavities in Swing Forging. In: Tobias, S.A. (Ed.), *Proceedings of the 17th International Machine and Tool Design Research Conference* (pp. 633-645), Birmingham, UK.

Cocks, A. (1989). Inelastic deformation of porous materials. *Journal of the Mechanics and Physics of Solids, 37*(6), 693-715. http://dx.doi.org/10.1016/0022-5096(89)90014-8

Dudra, S. P., & Im, Y. T. (1990). Analysis of void closure in open-die forging. *International Journal of Machine Tools & Manufacture, 30*(1), 65-75. http://dx.doi.org/10.1016/0890-6955(90)90042-H

Duva, J. M., & Hutchinson, J. W. (1984). Constitutive potentials for dilutely voided non-linear materials. *Mechanics of Materials, 3*(1), 41-54. http://dx.doi.org/10.1016/0167-6636(84)90013-9

Hill, R. (1956). New horizons in the mechanics of solids. *Journal of the Mechanics and Physics of Solids, 5*(1), 66-74. http://dx.doi.org/10.1016/0022-5096(56)90009-6

Jiang, Z., Ren, G. S., Xu, C. G., & Liu, G. H. (2005). An analyse simulation for single pore closure in hot cylinder ingots. *Journal of Plasticity Engineering, 12*(1), 47-49, 57.

Lee, B. J., & Mear, M. E. (1994). Studies of the growth and collapse of voids in viscous solids. *Journal of Engineering Materials and Technology, 116*(3), 348-358. http://dx.doi.org/10.1115/1.2904298

Lee, Y. S., Lee, S. U., Van Tyne, C. J., Joo, B. D., & Moon, Y. H. (2011). Internal void closure during the forging of large cast ingots using a simulation approach. *Journal of Materials Processing Technology, 211*(6), 1136-1145. http://dx.doi.org/10.1016/j.jmatprotec.2011.01.017

Park, C. Y., & Yang, D. Y. (1997). Modelling of void crushing for large-ingot hot forging. *Journal of Materials Processing Technology, 67*(1-3), 195-200. http://dx.doi.org/10.1016/S0924-0136(96)02843-9

Ståhlberg, U. (1986). Influence of spread and stress on the closure of a central longitudinal hole in the hot rolling of steel. *Journal of Mechanical Working Technology, 13*(1), 65-81.

Tanaka, M., Ono, S., & Tsuneno, M., (1986). Factors contributing to crushing of voids during forging. *J. Jpn. Soci. Technol. Plast., 27*(306), 927-934.

Zhang, X. X., Cui, Z. S., Chen, W., & Li, Y. (2009). A criterion for void closure in large ingots during hot forging. *Journal of Materials Processing Technology, 209*(4), 1950-1959.

Ranking Normalization Methods for Improving the Accuracy of SVM Algorithm by DEA Method

Maysam Eftekhary[1], Peyman Gholami[1], Saeed Safari[1] & Mohammad Shojaee[2]

[1] Department of Industrial Engineering, Arak Branch, Islamic Azad University, Arak, Iran

[2] Governmental Management, Arak, Iran

Correspondence: Peyman Gholami, Department of Industrial Engineering, Arak Branch, Islamic Azad University, Arak, Iran. Tel: 98-937-138-9131. E-mail: peyman711@gmail.com

Abstract

Data mining techniques, extracting patterns from large databases have become widespread in all life's aspect. One of the most important data mining tasks is classification. Classification is an important and widely studied topic in many disciplines, including statistics, artificial intelligent, operations research, computer science and data mining and knowledge discovery. One of the important things that should be done before using classification algorithms is preprocessing operations which cause to improve the accuracy of classification algorithms. Preprocessing operations include various methods that one of them is normalization. In this paper, we selected five applicable normalization methods and then we normalized selected data sets afterward we calculated the accuracy of classification algorithm before and after normalization. In this study the SVM algorithm was used in classification because this algorithm works based on n-dimension space and if the data sets become normalized the improvement of results will be expected. Eventually Data Envelopment Analysis (DEA) is used for ranking normalization methods. We have used four data sets in order to rank the normalization methods due to increase the accuracy then using DEA and AP-model outrank these methods.

Keywords: data mining, SVM, DEA, normalization methods, ranking, classification algorithm

1. Introduction

Data mining and knowledge discovery (DMKD) has made predominant progress during the past two decades (Peng et al., 2008). It utilizes methods, algorithms, and techniques from many disciplines, including statistics, datasets, machine learning, pattern recognition, artificial intelligence, data visualization, and optimization (Fayyad, 1996).

In recent years, the field of data mining has seen an explosion of interest from both academic and industry (e.g., Olafson, Li, & Wu, 2008). Increasing volume of data, increasing awareness of inadequacy of human brain to process data and increasing affordability of machine learning are reasons of growing popularity of data mining (e.g., Marakas, 2004). Data mining (DM) is the process for automatic discovery of high level knowledge by obtaining information from real data.

One of the major tasks in DMKD is classification. Researchers in a variety of fields have created a large number of classification algorithms, such as decision tree, neural networks, Bayesian network, linear logistic regression, Naive Bayes, and K-nearest-neighbor.

Learning algorithms are now used in many domains, and different performance metrics are suited for each domain. For example Precision/Recall measures are used in information retrieval; medicine prefers ROC area; Lift is appropriate for some marketing tasks, etc.

The different performance metrics measure different trade of in the predictions made by a classification and it is possible for learning methods to perform well on one metric, but be suboptimal on other metrics. Because of this it is to calculate algorithms on a broad set of performance metrics (Caruana & Niculescu-Mizil, 2006).

Classification is an important and widely studied topic in many disciplines, including statistics, artificial intelligent, operations research, computer science and data mining and knowledge discovery (Chen, Xu, & Chi, 1999).

Based on the number of predefined groups, classification can be divided into binary and multiclass classification. Binary classification assigns data objects into one of the two groups and multiclass classification involves three or more groups. Compare with binary classification, multiclass classification problem is more complex.

With increased use in real-world applications, such as disease diagnosis, text categorization, credit analysis, software risk management and network intrusion detection, a variety of methods and algorithms have been developed for multiclass classification in recent years.

Chen et al. scrutinized the reason why the mixture of experts (ME) performance poorly in multiclass classification and proposed an approximation for the Newton-Raphson algorithm to improve the performance of the ME architecture in multiclass classification.

Platt et al. (1999) presented the decision directed acyclic graph architecture and a learning algorithm for multiclass classification. Allwein et al. (2000) proposed a unifying framework for multiclass classification using a margin-based binary learning algorithm.

Crammer and Singer (2001) described the algorithmic implementation of multiclass kernel-based vector machines and conducted experiments to compare the presented approach to previously studied kernel-based methods. Loucopoulos proposed a mixed-integer programming model for the minimization of misclassification costs in the three-group problem.

Thorsten Joachims (1997) published results on a set of binary text classification experiments using the Support Vector Machine. The SVM concluded lower error than many other classification techniques.

Yang and Liu (1995) followed two years later with experiments of their own on the same data set. They used improved versions of Naive Bayes (NB) and k-nearest neighbors (KNN) but still found that the SVM performed at least as well as all other classifiers they tried.

Nowadays normalization methods are very applicable and appropriate for solving the Multiple Criteria Decision Making (MCDM) problems. There are different methods for normalization and these methods used to obtain concise answers. The normalization methods extracted from (Hwang & Yoon, 1981; Milani, Shanian, & Madoliat, 2005; Yoon & Hwang, 1995) and are used in the experimental study.

Ranking of normalization methods normally need to examine several criteria, such as accuracy, computational time, and misclassification rate. Therefore algorithm selection can be modeled as multiple criteria decision making (MCDM) problems (Peng et al., 2009; Rokach & Ensemble, 2009).

As mentioned heretofore, algorithm ranking is a useful strategy for selecting the appropriate classifier and the preferences of users are important in algorithm ranking (Berrer, Paterson, & Keller, 2000). Some existing MCDM methods are able to rank classifiers based on multiple performance measures and take the preferences of users into the ranking process.

DEA is a non-parametric linear programming based technique for measuring the relative efficiency of a set of similar units, usually referred to as decision making units (DMUs). Because of its successful application, DEA has gained too much attention and vast use by business and academic researchers.

Evaluation of data warehouse operations (e.g., Mannino, Hong, & Choi, 2008), selection of flexible manufacturing system (e.g., Camanho & Dyson, 2005) assessment of bank branch performance (e.g., Chen, Skully, & Brown, 2005) examining bank efficiency (e.g., Edirisinghe & Zhang, 2007), analyzing firm's financial statements (e.g., Johnes, 2006) measuring the efficiency of higher education (by solving only one LP) institutions (e.g., Ertay, Ruan, & Tuzkaya, 2006), solving facility layout design (FLD) problem (e.g., Shafer & Byrd, 2006) and measuring the efficiency of organizational investments in information technology (e.g., Vladimir Vapnik, 1995) are examples of using DEA in different areas.

The rest of this paper is organized as follows. In Section II normalization, SVM and DEA are described. The design of the experimental study is provided in Section III. Section IV presents and analyzes the experimental results. The Section V summarizes the findings and discusses future research directions.

2. Preliminaries

2.1 Normalization

The normalization methods and distance measures are also taken into consideration as well. Some of them are very useable that we selected 5 popular methods with study literature that such as:

1)   Vector normalization:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} \tag{1}$$

2)   Linear normalization (I):

$$r_{ij} = \frac{x_{ij}}{x_j^*} \tag{2}$$

$$i = 1,\ldots.m; j = 1,\ldots.n; x_j^* = \max\{x_{ij}\}$$

3)   Linear normalization (II):

$$r_{ij} = \frac{x_{ij} - x_j^{\sim}}{x_j^* - x_j^{\sim}} \tag{3}$$

$$i = 1,\ldots,m; j = 1,\ldots,n; x_j^{\sim} = \min\{x_{ij}\}$$

4)   Linear normalization (III):

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}} \tag{4}$$

$$i = 1,\ldots,m; j = 1,\ldots,n$$

5)   Non-monotonic normalization:

$$e^{\frac{-z^2}{2}}, z = \frac{(x_{ij} - x_j^0)}{\sigma_j} \tag{5}$$

x_j is the most favorable value and σ_j is the standard deviation of alternative ratings with respective to the jth attribute (Hwang & Yoon, 1981; Milani, Shanian, & Madoliat, 2005; Yoon & Hwang, 1995).

2.2 Support Vector Machine (SVM)

The Support Vector Machine is a classifier, originally proposed by Vapnik that finds a maximal margin separating hyper plane between two classes of data (e.g., Christopher, 1998). An SVM is trained via the following optimization problem:

$$Argmin_\omega \frac{1}{2}\omega^2 + C\sum_i \varepsilon_i \tag{6}$$

with constraints:

$$y_i\left(X_i.w + b\right) \geq 1 - \varepsilon_i \forall i \tag{7}$$

For more information, see Burges' tutorial and Cristianini and Shawe-Taylor's book (e.g., Nello & John, 2000; Yang & Liu, 1999). There are non-linear extensions to the SVM, but Yang and Liu found the linear kernel to outperform non-linear kernels in text classification (e.g., Ryan, 2000).

In informal experiments, we also found that linear performs at least as well as non-linear kernels. Hence, we only present linear SVM results. We use the SMART ltc[2] transform; the SvmFu package is used for running experiments (e.g., Charnes, Cooper, & Rhodes, 1978).

The SVM must read in the training set and then perform a quadratic optimization. This can be done quickly when the number of training examples is small (e.g., <10000 documents), but can be a bottleneck on larger training sets. We realize speed improvements with chunking and by caching kernel values between the training of binary classifiers.

2.3 Data Envelopment Analysis (DEA)

Charnes, Cooper, and Rhodes developed data envelopment analysis (DEA) to evaluate the efficiency of decision making units (DMUs) through identifying the efficiency frontier and comparing each DMU with the frontier (Koksalan & Tuncer, 2009). Since DEA is able to estimate efficiency with minimal prior assumptions (Li & Ma, 2008; Cherchye & Post, 2003), it has a comparative advantage to approaches that require a priori assumptions, such as standard forms of statistical regression analysis (Cooper, 2004).

During the past thirty years, various DEA extensions and models have been developed and established themselves as powerful analytical tools (Banker, Charnes, & Cooper, 1984).

The original DEA model presented by Charnes, Cooper, and Rhodes (Koksalan & Tuncer, 2009) is called ''CCR ratio model'', which uses the ratio of outputs to inputs to measure the efficiency of DMUs. Assume that there are n DMUs with m inputs to produce s outputs. X_{ij} and y_{rj} represent the amount of input *i* and output *r* for DMU_j (*j=1,2,....n*), respectively. Then the ratio-form of DEA can be represented as:

$$maxh_o(u.v) = \frac{\sum_r u_r y_{ro}}{\sum_i v_i x_{io}}$$

$$subject\ to : \sum_r u_r y_{ro} / \sum_i v_i x_{io} \leq 1\ for\ j = 1,\ldots,n \qquad (8)$$

$$u_r, v_i \geq 0\ for\ all\ i\ and\ r$$

where the u_r's and the v_i's are the variables and the y_{ro}'s and x_{io}'s are the observed output and input values of the DMU to be evaluated (i.e., DMU_o), respectively. The equivalent linear programming problem using the Charnes-Cooper transformation is (Banker, Charnes, & Cooper, 1984):

$$max\ z = \sum_{r=1}^{s} u_r y_{ro}$$

$$subject\ to : \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \qquad (9)$$

$$\sum_{i=1}^{m} v_i x_{ij} = 1$$

$$u_r, v_i \geq 0$$

Banker, Charnes, and Cooper introduced the BCC model by adding a constraint $\sum \lambda j = 1$ to the CCR model (Nakhaeizadeh & Schnabl, 1997). These models can be solved using the simplex method for each DMUs. DMUs with value of 1 are efficient and others are inefficient. Nakhaeizadeh and Schnabl proposed to use DEA approach in data mining algorithms selection (Fayyad, Piatetsky-Shapiro, & Smyth, 1996).

They argued that in order to make an objective evaluation of data mining algorithms that all the available positive and negative properties of algorithms are important and DEA models are able to take both aspects into consideration. Positive and negative properties of data mining algorithms can be considered as output and input components in DEA, respectively.

For example, the overall accuracy rate of a classification algorithm is an output component and the computation time of an algorithm is an input component. Using existing DEA models, it is possible to give a comprehensive evaluation of data mining algorithms.

In the empirical study, input components are 1. Output components include the accuracy SVM algorithm with use each normalization method and the DMU (Decision Making Unit) are the normalization method.

After solving this LP and determining the weights, the algorithms with $h_r = 1$ (100%) are efficient algorithms and form the efficiency frontier or envelope.

The other algorithms do not belong to the efficiency frontier and remain outside of it. As already mentioned, the definition of efficiency is more general than interestingness as suggested by Fayyad et al. (1993).

One can connect also both concepts in this form that more efficient algorithms are more interesting. For ranking the algorithms, one can use the approach suggested by Andersen and Petersen (2008). (AP-model). They use a criterion that we call it the AP-value.

In input-oriented models the AP-value measures how much an efficient algorithm can radically enlarge its input-levels while remaining still efficient (output-oriented is analogous). For example, for an input-oriented method an AP-value equal to 1.5 means that the algorithm remains still efficient when its input values are all enlarged by 50%. If the algorithm is inefficient then the AP-value is equal to the efficiency value. In this paper, the CCR and BCC models are utilized to rank a normalization method.

3. Experimental Study

The experiment is designed to rank normalization methods with the SVM classification algorithm using the DEA model described in the previous section. The following subsections describe the performance measures, data sources and experimental design.

3.1 Performance Measures

There are an extensive number of performance measures for classification. Commonly used performance measures in software defect classification are accuracy, precision, recall, F-measure, the area under receiver operating characteristic (AUC), and mean absolute error (Elish & Elish, 2008; Lessmann et al., 2008; Mair et al., 2000; Han & Kamber, 2006).

Besides these popular measures, this work includes seven other classification measures. The following paragraphs briefly describe these measures.

Overall accuracy: Accuracy is the percentage of correctly classified modules (Baeza-Yates & Ribeiro-Neto, 1999). It is one the most widely used classification performance metrics.

$$Overal\ accuracy = \frac{TN + TP}{TP + FP + FN + TN} \tag{10}$$

True positive (TP): TP is the number of correctly classified fault-prone modules. TP rate measures how well a classifier can recognize fault-prone modules. It is also called sensitivity measure.

$$\frac{True\ positive\ rate}{sensitivity} = \frac{TP}{TP + FN} \tag{11}$$

False positive (FP): FP is the number of non-fault-prone modules that is misclassified as fault-prone class. FP rate measures the percentage of non-fault-prone modules that were incorrectly classified.

$$False\ positive\ rate = \frac{FP}{FP + TN} \tag{12}$$

True negative (TN): TN is the number of correctly classified non-fault-prone modules. TN rate measures how well a classifier can recognize non-fault-prone modules. It is also called specificity measure.

$$\frac{True\ negtive\ rate}{specificity} = \frac{TN}{TN + FP} \tag{13}$$

False negative (FN): FN is the number of fault-prone modules that is misclassified as non-fault-prone class. FN rate measures the percentage of fault-prone modules that were incorrectly classified.

$$False\ negative\ rate = \frac{FN}{FN + TP} \tag{14}$$

Precision: This is the number of classified fault-prone modules that actually are fault-prone modules.

$$Pr\ ecision = \frac{TP}{TP + FP} \tag{15}$$

Recall: This is the percentage of fault-prone modules that are correctly classified.

$$Re\ call = \frac{TP}{TP + FN} \tag{16}$$

F-measure: It is the harmonic mean of precision and recall. F-measure has been widely used in information retrieval (Ferri, Hernandezorallo, & Modroiu, 2009).

$$F - measure = \frac{2 \times Pr\ ecision \times Re\ call}{Pr\ ecision + Re\ call} \tag{17}$$

AUC: ROC stands for Receiver Operating Characteristic, which shows the tradeoff between TP rate and FP rate

(Ferri, Hernandezorallo, & Modroiu, 2009). AUC represents the accuracy of a classifier. The larger is the area, the better is the classifier.

Kappa statistic (KapS): This is a classifier performance measure that estimates the similarity between the members of an ensemble in multi-classifiers systems (Witten & Frank, 2005).

$$KapS = \frac{P(A) - P(E)}{1 - P(E)} \tag{18}$$

P(A) is the accuracy of the classifier and P(E) is the probability that agreement among classifiers is due to chance.

$$P(E) = \frac{\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{c}\sum_{i=1}^{m}f(i,k)C(i,j)\right]\left[\sum_{j=1}^{c}\sum_{i=1}^{m}f(i,f)C(i,k)\right]\right)}{m'} \tag{19}$$

m is the number of modules and c is the number of classes. $f(i, j)$ is the actual probability of i module to be of

class $\sum_{i=1}^{m} f(i,j)$ is the number of modules of class j. Given threshold θ, $C_\theta (i, j)$ is 1 if j is the predicted class

for i obtained from $P(i, j)$; otherwise it is [0, 1] (Triantaphyllou & Baig, 2005).

Mean absolute error (MAE): This measures how much the predictions deviate from the true probability. $P(i, j)$ is the estimated probability of i module to be of class j taking values in [0,1] (Triantaphyllou & Baig, 2005).

$$MAE = \frac{\sum_{j=1}^{C}\sum_{j=1}^{m}|f(i,j)]) - p(i,j)|}{m.c} \tag{20}$$

As the name suggests, the mean absolute error is an average of the absolute errors $e_i = f_i - y_i$, where f_i is the prediction and y_i the true value. Note that alternative formulations may include relative frequencies as weight factors. The mean absolute error is one of a number of ways of comparing forecasts with their eventual outcomes.

Well-established alternatives are the mean absolute scaled error (MASE) and the mean squared error. These all summarize performance in ways that disregard the direction of over- or under- prediction; a measure that does place emphasis on this is the mean signed difference.

Where a prediction model is to be fitted using a selected performance measure, in the sense that the least squares approach is related to the mean squared error, the equivalent for mean absolute error is least absolute deviations.

- Training time: the time needed to train a classification algorithm or ensemble method.
- Test time: the time needed to test a classification algorithm or ensemble method.

3.2 Data Sources

The data used in this study are 6 public-domain data sets from four application domains including Iris, Glass and Breast Cancer and Breast Cancer Wisconsin are provided by the UCI machine learning repository (http://archive.ics.uci.edu/ml/).

The Iris is a three-class (Iris Setosa, Iris Versicolour, Iris Virginica) data set that has 150 instances with 4 continuous predictor variables and 1 class variable. The predictor variables describe the Sepal length, Sepal width, Petal length and Petal width of Iris plants and the class variable indicates the type of iris plant.

The Breast Cancer is a two-class (no-recurrence-events, recurrence-events) dataset that has 699 instances with 9 continuous predictor variables and 1 class variable. The predictor variables describe the Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli, mitoses.

The Breast Cancer Wisconsin is a two-class(R = recur, N = nonrecur) dataset that has 198 instances with 10 continuous predictor variables and 1 class variable. The predictor variables describe the radius (mean of distances from center to points on the perimeter), texture (standard deviation of gray-scale values), perimeter, area, smoothness (local variation in radius lengths), compactness (perimeter^2/area - 1.0), concavity (severity of concave portions of the contour), concave points (number of concave portions of the contour),
symmetry, fractal dimension ("coastline approximation" - 1).

The Cardiotocography is a three-class (N=normal; S=suspect; P=pathologic) dataset that has 15888 instances

with 22 continuous predictor variables and 1 class variable. The predictor variables describe the LB - FHR baseline (beats per minute), AC-#of accelerations per second, FM - # of fetal movements per second, UC - # of uterine contractions per second, DL - # of light decelerations per second, DS - # of severe decelerations per second, DP - # of prolonged decelerations per second ASTV - percentage of time with abnormal short term variability, MSTV - mean value of short term variability ALTV - percentage of time with abnormal long term variability, MLTV - mean value of long term variability, Width-width of FHR histogram, Min-minimum of FHR histogram, Max - Maximum of FHR histogram, Nmax - # of histogram peaks, Nzeros - # of histogram zeros, Mode - histogram mode, Mean - histogram mean, Median - histogram median, Variance - histogram variance, Tendency - histogram tendency, CLASS - FHR pattern class code (1 to 10).

The Wall-Following Robot Navigation is a four-class (Move-Forward, Slight-Right-Turn, Sharp-Right-Turn, Slight-Left-Turn) data set that has 5456 instances with 24 continuous predictor variables and 1 class variable. The predictor variables describe US1: ultrasound sensor at the front of the robot (reference angle: 180°), US2: ultrasound reading (reference angle: -165°), US3: ultrasound reading (reference angle: -150°), US4: ultrasound reading (reference angle: -135°), US5: ultrasound reading (reference angle: -120°), US6: ultrasound reading (reference angle: -105°), US7: ultrasound reading (reference angle: -90°), US8: ultrasound reading (reference angle: -75°), US9: ultrasound reading (reference angle: -60°), US10: ultrasound reading (reference angle: -45°), US11: ultrasound reading (reference angle: -30°), US12: ultrasound reading (reference angle: -15°), US13: reading of ultrasound sensor situated at the back of the robot (reference angle: 0°), US14: ultrasound reading (reference angle: 15°), US15: ultrasound reading (reference angle: 30°), US16: ultrasound reading (reference angle: 45°), US17: ultrasound reading (reference angle: 60°), US18: ultrasound reading (reference angle: 75°), US19: ultrasound reading (reference angle: 90°), US20: ultrasound reading (reference angle: 105°), US21: ultrasound reading (reference angle: 120°), US22: ultrasound reading (reference angle: 135°), US23: ultrasound reading (reference angle: 150°), US24: ultrasound reading (reference angle: 165°).

3.3 Experimental Design

The experiment was carried out according to the following process:

Input: the four selected datasets.

Output: Ranking of normalization techniques.

Step 1: Prepare target datasets: select and transform relevant features; data cleaning; data integration.

Step 2: Train and test the SVM classification model on a randomly sampled partitions (i.e., 10-fold cross-validation) using WEKA 3.7 (Shih, Huan-JyhShyur, & Lee, 2007).

Step 3: Normalized the data set by five methods.

Step 4: Train and test the SVM classification model on a randomly sampled partitions after normalization the data sets (i.e., 10-fold cross-validation) using WEKA 3.7 (Shih, Huan-JyhShyur, & Lee, 2007).

Step 5: Ranking the Normalization methods with attention accuracies get from SVM classification algorithm by DEA method.

Step 6: If gained efficiency from DEA model of two or more algorithms equal 1 then ranking Normalization methods with using A-P model of DEA.

END

Normalization methods selection problem involves benefit criteria. A criterion is called the benefit because the higher a normalization methods scores in terms of the corresponding criterion, the better the algorithm is (Abbas, Morteza, & Karimi, 2011).

4. Discussion of Results

Table 1 shows the classification results of 4 data sets using SVM classification algorithm before normalization methods.

Table 1. Classification results before normalization methods

Accuracy(SVM)	Data sets
Iris	96
Glass	80.84
Breast Cancer	96.71
Breast Cancer Wisconsin	96.71
Cardiotocography	60.35
Wall-Following Robot Navigation	74.8

Table 2 shows the classification results of 4 data sets using SVM classification algorithm after normalization methods.

Table 2. Classification results after normalization methods

Data set	Vector normalization	Accuracy of normalization Methods			
		Linear normalization (I)	Linear normalization (II)	Linear normalization (III)	Non-monotonic normalization
Iris	96.71	96.27	96	96	97.07
Glass	81	81	80.84	80.84	81.97
Breast Cancer	97.14	96.71	96.6	96.6	97.02
Breast Cancer Wisconsin	97.71	97.05	97.05	98.27	98.82
Cardiotocography	63.06	61	64.2	62.09	68.05
Wall-Following Robot Navigation	77.52	76.02	74.92	75.81	77.3

Table 3 represents the ranking results generated by DEA. That the efficiency of all normalization methods is 1 it means all normalization methods are efficient and for ranking the normalization methods we should rank them by A.P model.

Table 3. Ranking results generated by DEA

Normalization methods	Efficiency
Vector normalization	1
Linear normalization (I)	1
Linear normalization (II)	1
Linear normalization (III)	1
Non-monotonic normalization	1

Table 4 represents the ranking results generated by A.P model and ultimate evaluation of them.

Table 4. Represents the ranking results generated by A.P

Normalization methods	Overall efficiency by A.P Model	Final Ranking
Vector normalization	6.71	2
Linear normalization (I)	4.26	4
Linear normalization (II)	3.11	5
Linear normalization (III)	5.10	3
Non-monotonic normalization	8.42	1

5. Conclusions and Future Work

In this paper we implemented different normalization methods as a preprocessing to improve SVM algorithm. After testing the accuracy of SVM method on various Data Sets we concluded that these simple methods have low amount of processing and also increase the accuracy of SVM algorithm. For ranking these methods we used DEA which is the most practical decision making techniques.

We suggest to the future researchers to apply normalization methods on more data sets and also use other decision making techniques like TOPSIS and VIKOR (Peng et al., 1999) for ranking them. In addition testing these methods on other classification algorithms especially those based on n-dimension space like SMO to see how their accuracy improves is also suggested. And also they considerate time complexity in their researches.

References

Allwein, E. L., Schapire, R. E., Singer, Y., & Kaelbling, P. (2000). Reducing multiclass to binary: a unifying approach for margin classifiers. *Journal of Machine Learning Research, 1*, 113-124.

Andersen, P., & Petersen, N. C. (1993). A Procedure for Ranking Efficient Units in Data Envelopment Analysis. *Management Science, 39*(10), 1261-1264. http://dx.doi.org/10.1287/mnsc.39.10.1261

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison, Wesley. New York, NY: ACM Press.

Banker, R., Charnes, A., & Cooper. W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*, 1078-1092. http://dx.doi.org/10.1287/mnsc.30.9.1078

Bazzazi, A. A., Osanloo, M., & Karimi, B. (2011). Deriving preference order of open pit mines equipment through MADM methods: Application of modified VIKORE method. *Expert Systems with Applications, 38*, 2550-2556. http://dx.doi.org/10.1016/j.eswa.2010.08.043

Berrer, H., Paterson, I., & Keller, J. (2000). Evaluation of machine-learning algorithm ranking advisors. in: P. Brazdil, A. Jorge (Eds.), Proceedings of the PKDD2000Workshop on Data Mining, Decision Support, Meta-Learning and ILP: Forum for Practical Problem Presentation and Prospective Solutions, 1-13.

Camanho, A. S., & Dyson, R. G. (2005). Cost efficiency measurement with price uncertainty: A DEA application to bank branch assessments. *European Journal of Operational Research, 161*, 432-446. http://dx.doi.org/10.1016/j.ejor.2003.07.018

Caruana, R., & Niculescu-Mizil, A. (2006). An Empirical Comparison of Supervised Learning Algorithms. Appearing in Proceedings of the 23 rd International Conference on Machine Learning, Pittsburgh, PA.

Challagulla, V. U. B., Bastani, F. B., Raymond, I. Y., & Paul, A. (2008). Empirical assessment of machine learning based software defect prediction techniques. *International Journal on Artificial Intelligence Tools, 17*(2), 389-400. http://dx.doi.org/10.1142/S0218213008003947

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444. http://dx.doi.org/10.1016/0377-2217(78)90138-8

Chen, K., Xu, L., & Chi, H. (1999). Improved learning algorithms for mixture of experts in multiclass classification. *Neural Networks, 12*, 1229-1252. http://dx.doi.org/10.1016/S0893-6080(99)00043-X

Chen, X., Skully, M., & Brown, K. (2005). Banking efficiency in China: Application of DEA to pre- and post-deregulation eras: 1993-2000. *China Economic Review, 16*, 229-245. http://dx.doi.org/10.1016/j.chieco.2005.02.001

Cherchye, L., & Post, T. (2003). Methodological advances in DEA: a survey and an application for the Dutch electricity sector. *StatisticaNeerlandica, 57*(4), 410-438. http://dx.doi.org/10.1111/1467-9574.00238

Christopher, J. C. B. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery, 2*(2), 121-167. http://dx.doi.org/10.1023/A:1009715923555

Cooper, W. W., Seiford, L. M., & Zhu, J. (2004). Data envelopment analysis: history, models and interpretations. *Handbook on Data Envelopment Analysis.* Boston: Kluwer Academic Publisher, Chapter 1, 1-39.

Crammer, K., & Singer, Y. (2001). On the algorithmic implementation of multiclass kernel based vector machines. *Journal of MachineLearning Research, 2*, 265-292.

Edirisinghe, N. C. P., & Zhang, X. (2007). Generalized DEA model of fundamental analysis and its application to portfolio optimization. *Journal of Banking & Finance, 31*, 3311-3335. http://dx.doi.org/10.1016/j.jbankfin.2007.04.008

Elish, K. O. (2008). Predicting defect-prone software modules using support vector machines. *Journal of Systems and Software, 81*(5), 649-660. http://dx.doi.org/10.1016/j.jss.2007.07.040

Ertay, T., Ruan, D., & Tuzkaya, U. R. (2006). Integrating data envelopment analysis and analytic hierarchy for the facility layout design in manufacturing systems. *Information Sciences, 176*, 237-262. http://dx.doi.org/10.1016/j.ins.2004.12.001

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery: an overview. *Advances in Knowledge Discovery and Data Mining,* 1-34.

Ferri, C., Hernandezorallo, J., & Modroiu, R. (2009). An experimental comparison of performance measures for classification. *Pattern Recognition Letters,* 27-38. http://dx.doi.org/10.1016/j.patrec.2008.08.010

Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques (2nd ed.).* San Francisco: Morgan Kaufman.

Hsu-Shih, S., Huan-Jyh, S., & Lee, E. S. (2007). An extension of TOPSIS for group decision making. *Mathematical and Computer Modeling, 45*, 801-813. http://dx.doi.org/10.1016/j.mcm.2006.03.023

Hwang, C. L., & Yoon, K. (1981). *Multiple Attribute Decision Making.* Berlin: Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-48318-9

Johnes, J. (2006). Measuring teaching efficiency in higher education: An applicationof data envelopment analysis to economics graduates from UK Universities. *European Journal of Operational Research, 174*, 443-456. http://dx.doi.org/10.1016/j.ejor.2005.02.044

Koksalan, M., & Tuncer, C. (2009). A DEA-based approach to ranking multi-criteria alternatives. *International Journal of Information Technology and Decision Making, 8*(1), 29-54. http://dx.doi.org/10.1142/S0219622009003259

Kuncheva, L. I. (2004). *Combining Pattern Classifiers: Methods and Algorithms.* Wiley. http://dx.doi.org/10.1002/0471660264

Lessmann, S., Baesens, B., Mues, C., & Pietsch, S. (2008). Benchmarking classification models for software defect prediction: a proposed framework and novel findings. *IEEE Transactions on Software Engineering, 34*(4), 485-496. http://dx.doi.org/10.1109/TSE.2008.35

Li, H., & Ma, L. (2008). Ranking decision alternatives by integrated DEA, AHP and gower plot techniques. *International Journal of Information Technology & Decision Making, 7*(2), 241-258. http://dx.doi.org/10.1142/S0219622008002922

Liu, S. T. (2008). A fuzzy DEA/AR approach to the selection of flexible manufacturing systems. *Computers & Industrial Engineering, 54*, 66-76. http://dx.doi.org/10.1016/j.dss.2007.10.011

Loucopoulos, C. (2001). Three-group classification with unequalMisclassification costs: a mathematical programming approach. *Omega, 29*(3), 291-297. http://dx.doi.org/10.1016/S0305-0483(01)00023-8

Mair, C., Kadoda, G., Leflel, M., Phapl, L., Schofield, K., Shepperd, M., & Webster, S. (2000). An investigation of machine learning based prediction systems. *Journal of Systems Software, 53*(1), 23-29. http://dx.doi.org/10.1016/S0164-1212(00)00005-4

Mannino, M., Hong, S. N., & Choi, I. J. (2008). Efficiency evaluation of data warehouse operations. *Decision Support Systems, 44*, 883-898. http://dx.doi.org/10.1016/j.dss.2007.10.011

Marakas, G, M. (2004). *Decision support system in the 21ˢᵗ Centry (2ⁿᵈ ed.)*. Prentice-Hall of India.

Milani, A. S., Shanian, A., & Madoliat, R. (2005). The effect of normalization norms in multiple attribute decision making model: A case study in gear material selection. *Structure Multiciplinary Optimization, 29*(4), 312-318.

Nakhaeizadeh, G., & Schnabl, A. (1997). Development of multi-criteria metrics for evaluation of data mining algorithms. In Proceeding of the Third International Conference on Knowledge Discovery and Data Mining (KDD'97), Newport Beach, California, August 14-17, 37-42.

Nello, C., & Shawe-Taylor, J. (2000). *An introduction to support vector machines.* Cambridge University Press.

Olafson, S., Li, X., & Wu, S. (2008). Operations research and data mining. *European Journal of Operational Research, 187*, 1429-1448. http://dx.doi.org/10.1016/j.ejor.2006.09.023

Peng, Y., Kou, G., Shi, Y., & Chen, Z. X. (2008). A descriptive framework for the field of data mining and knowledge discovery. *International Journal of Information Technology and Decision Making, 7*(4), 639-682. http://dx.doi.org/10.1142/S0219622008003204

Peng, Y., Kou, G., Wang, G., Wang, H., & Ko, F. (2009). Empirical evaluation of classifiers for software risk management. *International Journal of Information Technology and Decision Making, 8*(4), 749-768. http://dx.doi.org/10.1142/S0219622009003715

Platt, J. C., Cristianini, N., Shawe-taylor, J. (1999). Large margin DAGs for multiclass classification. In Proceedings of neural information processing systems, NIPS'99, MIT Press. 547-553.

Rifkin, R. (2000). Svmfu. Retrieved from http://five-percent-nation.mit.edu/SvmFu/

Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review, 33*(1-2), 1-39. http://dx.doi.org/10.1007/s10462-009-9124-7

Shafer, S. M., & Byrd, T. A. (2006). A framework for measuring the efficiency of organizational investments in information technology using data envelopment analysis. *Omega, 28*, 125-141. http://dx.doi.org/10.1016/S0305-0483(99)00039-0

Thorsten, J. (1997). *Text categorization with support vector machines: Learning with many relevant features.* Technical report, University of Dortmund, Computer Science Department.

Triantaphyllou, E., & Baig, K. (2005). The impact of aggregating benefit and cost criteria in four MCDA methods. *IEEE Transactions on Engineering Management, 52*, 213-262. http://dx.doi.org/10.1109/TEM.2005.845221

Vladimir, V. (1995). *The Nature of Statistical Learning Theory* (2ⁿᵈ ed.). New York, NY: Springer- Verlag.

Witten, I. H., & Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques *(2ⁿᵈ ed.)*. San Francisco: Morgan Kaufman.

Yang, Y. M., & Liu, X. (1995). *A re-examination of text categorization methods.* In Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval.

Yang, Y. M., & Liu, X. (1999). *A re-examination of text categorization methods.* In Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval.

Yoon, K. P., & Hwang, C. L. (1995). *Multiple Attribute Decision Making: An Introduction.* London: Sage Pub.

A Terminological Search Algorithm for Ontology Matching

Ahmad Zaeri[1] & Mohammad Ali Nematbakhsh[1]

[1] Computer Department, University of Isfahan, Isfahan, Iran

Correspondence: Ahmad Zaeri, Computer Department, University of Isfahan, Hezar Jarib Street, Isfahan, Iran. Tel: 98-311-793-4010. E-mail: zaeri@eng.ui.ac.ir

Abstract

Most of the ontology alignment tools use terminological techniques as the initial step and then apply the structural techniques to refine the results. Since each terminological similarity measure considers some features of similarity, ontology alignment systems require exploiting different measures. While a great deal of effort has been devoted to developing various terminological similarity measures and also developing various ontology alignment systems, little attention has been paid to develop similarity search algorithms which exploit different similarity measures in order to gain benefits and avoid limitations. We propose a novel terminological search algorithm which tries to find an entity similar to an input search string in a given ontology. This algorithm extends the search string by creating a matrix from its synonym and hypernyms. The algorithm employs and combines different kind of similarity measures in different situations to achieve a higher performance, accuracy, and stability in comparison with previous methods which either use one measure or combine more measures in a naive ways such as averaging. We evaluated the algorithm using a subset of OAEI Bench mark data set. Results showed the superiority of proposed algorithm and effectiveness of different applied techniques such as word sense disambiguation and semantic filtering mechanism.

Keywords: terminological search, similarity measures, semantic similarity, ontologies matching

1. Introduction

The vision of Semantic Web is about having data and knowledge machine understandable so that machines can analyze and process complex request of humans more efficiently. In other words, the semantic web should facilitate information sharing in any form and integrate information from different sources such as web contents or database records (Berners-Lee, Hendler, & Lassila, 2001). An initial step toward this vision has been taken by representing terminologies of different domains as Ontologies. Even in the same domain, having different ontologies cannot be avoided due to complexity or expansiveness of knowledge or contrasting and distinctive views of different users (Romero, Vázquez-Naya, Loureiro, & Ezquerra, 2009). Consequently, to successfully integrating data sources with different ontologies, it is needed to align their ontologies through a process called Ontology Alignment. In its simplest form, an ontology alignment is finding of one to one correspondences among entities of two ontologies. In recent years, a large number of ontology alignment systems have been developed to detect such correspondences.

These systems usually employ different alignment techniques which have been categorized by Euzenat et al. (2007) in four main classes: terminological, structural, extensional, and semantic techniques. Terminological techniques try to find correspondences by investigating similarities between entities names. Alternatively, the structural methods consider internal structure of an entity or its relations to other entities as a source of detecting correspondences. Most of the developed alignment tools use terminological techniques as the initial and the main alignment approach and then apply the structural techniques to refine the results and to improve the accuracy. Besides these two main techniques, some systems use extensional and semantic techniques. Extensional techniques are inspired by the idea that having more commonalities between two entities' individuals (i.e., instances) might imply higher probability of matching between them. Although semantic techniques usually employ theoretical models and deduction to find similarity between interpretations of entities, the inductive nature of ontology alignment makes using of such deductive techniques difficult. Thus, semantic techniques are mostly utilized for validation of detected correspondences (Euzenat & Shvaiko, 2007).

Terminological techniques are divided into two main groups, the *string based* and the *language based*

approaches. The first one are techniques that just consider entity names as a sequence of characters and assume that higher similarity between structures of two sequences shows a higher similarity between them. For example, these techniques consider *PublishedBy* highly similar to *Publisher*, whereas they distinguish *Paper* from *Article*. The second one considers occurrence of some meaningful words in the names and compare two names by similarity between the meaning of those words (e.g., by using some external thesaurus). For example, these approaches can easily detect similarity between *Paper* and *Article*.

However, each of these measures has its own cons and pros. For example, string based measures usually have higher computational speeds and are more immune to small morphological changes (e.g., Food and Foods), but they are more sensitive to none morphological changes (e.g., *Food* and *Flood*). On the other hand, language based measures are less sensitive to none morphological changes (e.g., difference between *Food* and *Flood* can be easily detected), but they are much slower and are more sensitive to morphological changes. Although stemmer algorithms could help to find the root of a word, it would not be so beneficial if morphological changes highly alter the meaning of the word (e.g., *rewriter* and *writing*).

Each terminological similarity measure considers some aspects or features of similarity, so ontology alignment systems require exploiting different measures to achieve higher accuracy. Based on using similarity measures, ontology alignment systems can be divided into two groups: First, alignment systems which directly use similarity measures or combine different measures by some special technique such as averaging or weighted averaging; Second, alignment systems which independently developed their own techniques without using well known terminological measures.

While a great deal of effort has been devoted to developing various terminological similarity measures and also developing various ontology alignment systems, little attention has been paid to develop similarity search algorithms which exploit different similarity measures in order to gain their benefits and avoid their limitations.

In this paper, we suggest a novel terminological similarity search algorithm to find a concept Name (property or individual) in an ontology that is similar to a search string *SimString*. This algorithm extracts all synonyms and hypernyms of *SimString* from the WordNet and then creates a matrix which each row of this matrix represents one meaning of *SimString*. In fact, this algorithm not only tries to find concepts similar to *SimString* but also tries to find concepts similar to synonyms and hypernyms of all different *SimString* meanings as well. In addition, this algorithm prioritizes the synonyms and hypernyms in which it considers more specific words firstly then tries more general ones, so it can handle the situations that besides equal concepts, super concepts are interested. Once all ontology concepts names were compared to matrix candidates, each row of matrix, which represents one meaning of *SimString*, is scanned to find the at most one candidate per each row. This algorithm applies special kind of semantic filtering by removing all selected candidates that have a semantic similarity to *SimString* less than a threshold. The JIANG similarity measure has been used in the filtering part. In other words, we suppose that having a high semantic similarity is necessary but not sufficient to consider two words similar. ISub (Stoilos, Stamou, & Kollias, 2005) lexical similarity measure, as reported, has a clear superiority in ontology name searching over other measures such as Levenstein. Hence, ISub has been used in calculating of syntactical similarity calculation. Finally, algorithm uses a word sense disambiguation based on averaging to select the final similar concept (property or individual). This algorithm employs and combines different kind of similarity measures in different situations to achieve a higher performance, accuracy, and stability in comparison with either using one measure separately or combining different measures in simple ways such as averaging.

The rest of this article is structured as follows. After giving an overview of related work in Section 2, we illustrate the main algorithm in Section 3. Following this, Section 4 provides details on the experiments that we carried out and the results achieved. In these experiments we showed the superiority of our algorithm by using bench mark data set (Euzenat, Meilicke, Stuckenschmidt, Shvaiko, & dos Santos, 2011). We conclude with discussion of the approach and topics for future work in Section 5.

2. Related Work

Today, large numbers of ontology alignment systems exist in literature. Most of these alignment systems use terminological techniques as the first step to find the correspondences between ontologies' entities. These techniques usually utilize similarity measures to find such correspondences. Different similarity measures have been proposed to exploit different aspects of similarities between entities. As mentioned earlier, in literature, terminological similarity measures are usually divided into *string based* and *language based* measures.

Most widely used string based measure is the *Levenshtein distance* (Levenshtein, 1966). The Levenshtein distance between two strings is the minimum number of insertion, deletion, or substitution of a single character needed to transform one string into the other. Commonly used in bioinformatics, *Needleman-Wunsch distance*

(Needleman & Wunsch, 1970) is the modified version of Levenshtein distance which applies higher weight to insert and delete operations in comparison with substitution. *Jaro distance* (1995) and its modified variant *Jaro-Winkler distance* (1999) have been proposed for matching the names that may contain some spelling errors. Jaro measure considers common character between two strings and their order to calculate the similarity. Jaro measure is defined as follow (Euzenat & Shvaiko, 2007):

$$\text{Jaro}(s_1, s_2) = \frac{1}{3} \times \left(\frac{|\text{com}(s_1, s_2)|}{|s_1|} + \frac{|\text{com}(s_2, s_1)|}{|s_2|} + \frac{|\text{com}(s_1, s_2)| - |\text{transp}(s_1, s_2)|}{|\text{com}(s_1, s_2)|} \right) \tag{1}$$

which *com(s₁,s₂)* is the common character between *s₁* and *s₂* and *transp(s₁,s₂)* is the characters in *com(s₁, s₂)* with different orders in *s₁* and *s₂*.

Another feature for calculating similarity is to consider the number of common substrings between two strings.

$$subsim(s_1, s_2) = \frac{2 \times |maxCommonSubstring(s_1, s_2)|}{|s_1| + |s_2|} \tag{2}$$

Stoilos et al. (2005) have proposed ISUB measure which has specially been designed for ontology alignment by extending the idea of subsim measure:

$$ISUB(s_1, s_2) = Comm(s_1, s_2) - Diff(s_1, s_2) + winkler(s_1, s_2) \tag{3}$$

which

$$Comm(s_1, s_2) = \frac{2 \times \sum_i |maxComSubString_i|}{|s_1| + |s_2|} \tag{4}$$

and *maxCommonSubstringᵢ* extends the idea of maximum common substring by considering the next common substrings after removing previous common substrings.

$$Diff(s_1, s_2) = \frac{uLen(s_1) * uLen(s_2)}{0.6 + 0.4 \times (uLen(s_1) + uLen(s_2) - uLen(s_1) * uLen(s_2))} \tag{5}$$

uLen represents the length of the unmatched substring from the initial strings. $winkler(s_1, s_2)$ is the Jaro-Winkler similarity measure added for extra improvement.

They argue that for the case of ontology matching ISUB has a higher performance in comparison with other string based measures in the term of F1, precision, and recall (Stoilos et al., 2005).

In contrast to *string based measures*, *language based measures* consider string in the word level other than character level. In literature, language based measures are included two main groups: *intrinsic* and *extrinsic*. Intrinsic measures employ some linguistics techniques such as stemming, removing stop words, and part of speech tagging to find similarities between words while extrinsic measures uses some external resources such as dictionary and thesaurus to match words by considering their meaning. Many of extrinsic measures in ontology world utilize WordNet (Fellbaum, 1998) as an external resource to calculate extrinsic similarity measures (Euzenat & Shvaiko, 2007).

The WordNet extrinsic measures are divided into the three categories based on the kind of information that they consider: *path based*, *information content based*, and *hybrid measures*.

Path based measures: These measures use distance between two words' node in the taxonomy graph and their place to calculate the similarity. Higher distance between two nodes shows lower similarity between words.

Rada et al. (Rada, Mili, Bicknell, & Blettner, 1989) uses the length of path between two concepts as the distance between two concepts.

$$radaDistance = length(path(C_1, C_2)) \tag{6}$$

Leacock and Chodorow (1998) have normalized the Rada distance by using a *D* factor as the depth of the taxonomy contains the two concepts. Then they have translated it to a similarity measures as follow:

$$LCSim(C_1, C_2) = -log\left(\frac{length(path_{min}(C_1, C_2))}{2 \times D} \right) \tag{7}$$

Wu and Palmer (1994) define the similarity of two concepts based on their distance to the lowest common super concept and the distance of the common super concept to the root of taxonomy as well. The basic idea here is

that as the common subsumer goes far from the root the similarity of two concepts is become less sensitive to the distance of two concepts.

$$WuPalmerSim(C_1, C_2) = \frac{2 \times N_3}{N_1 + N_2 + 2 \times N_3} \qquad (8)$$

where C_3 is common subsumer of C_1 and C_2. If C_1 and C_2 have more than one common subsumer, the C_3 would be the most specific one. N_1, N_2, and N_3 are the lengths of the paths between C_1 to C_3, C_2 to C_3, and C_3 to *root* respectively.

Information content based measures (Cross & Hu, 2011): Finding path between two words in a taxonomy graph is usually a time consuming task, so information content based measures just employ the content of two nodes to determine the similarity between their corresponding words. Taxonomy is enriched by an information content function as follows:

$$IC(c) = -log(p(c)) \qquad (9)$$

IC represents the probability of encountering an instance of a concept c to all other concept instances.

Resnik (1995) have introduced the first measure based on the information content function.

$$ResnikSim(C_1, C_2) = max_{c \in S(C_1, C_2)}[IC(c)] \qquad (10)$$

where $S(C_1, C_2)$ is the set of C_1 and C_2 common subsumers. This measure just considers the common subsumer which has the highest amount of information content.

Lin (1998) extends Resnik measure to also consider information content of C_1 and C_2 as well.

$$LinSim(C_1, C_2) = \frac{2 \times IC(C_3)}{IC(C_1) + IC(C_2)} \qquad (11)$$

Jiang and Conrath (1997) have proposed another distance measure based on information content, but this measure has inspired by different idea. More data on common subsumer rather than on nodes themselves, the higher of the probability of similarity between the nodes.

$$JiangDistance(C_1, C_2) = IC(C_1) + IC(C_2) - 2 \times IC(C_3) \qquad (12)$$

where C_3 is the common subsumer of C_1 and C_2 with the highest information content.

Combined measures: These measures utilize combinations of different measures. For example, it could exploit the positions of two nodes in taxonomy graph as well as their contents to find the similarity between two concepts.

Pirro (2009) has combined the idea of feature based similarity with the information content based measures to propose a new measure. Feature based similarity have been suggested by Tversky et al. (1977) employs the common features of C_1 and C_2 as well as their differentiating features as follows:

$$TverskySim(C_1, C_2) = \frac{| fe(C_1) \cap fe(C_2) |}{| fe(C_1) \cap fe(C_2) | + \alpha | fe(C_1)/fe(C_2) | + \beta | fe(C_2)/fe(C_1) |} \qquad (13)$$

which fe is the feature set of the concept, and $\alpha, \beta \geq 0$ are the parameters of the Tversky similarity.

The Pirro similarity measure is defined as (Pirro, 2009):

$$PirroSim(C_1, C_2) = 3 \times IC(C_3) - IC(C_1) - IC(C_2) \qquad (14)$$

where C_3 is the most informative common subsumer as defined in Resnik measure.

After briefly reviewing terminological similarity measures, we now discuss the terminological matching techniques used in ontology matching systems. Some of these ontology matching systems directly employ those similarity measures while some others use special techniques to perform the terminological matching.

OLA (Euzenat & Valtchev, 2003) utilizes a measure derived from Wu-Palmer for terminological mapping. ASMOV (Jean-Mary, Shironoshita, & Kabuka, 2009) uses Wu-Palmer to find similarity between properties while uses Lin measure to find similarity between concept names.

Having one of the highest results in the OAEI contest, Agreement Maker (Cruz, Antonelli, & Stroe, 2008) uses three different matchers for terminological matching. These matchers are Base Similarity Matcher (BSM), Parametric String-based Matcher (PSM), and the Vector-based Multi-word Matcher (VMM).

Agreement Maker uses BSM to generate the initial alignments among concept names of two ontologies. BSM first tokenizes the entity compound names and then removes all stop words (such as, "the", "a"). It then uses the WordNet to enrich each word in the tokenized string by its glossary. BSM then employs the stemming algorithm to translate the words of enriched strings to their roots. After these preprocessing steps, BSM uses following similarity measure to calculate the similarity between the two enriched concept names:

$$BaseSim(C_1, C_2) = \frac{2 \times |D \cap D'|}{|D| + |D'|} \tag{15}$$

where D and D' are enriched versions of C_1 and C_2 respectively.

In PSM, users can choose various parameters which are suitable to the specific application. Users can select a set of string similarity measures such as Levenshtein or Jaro-Winkler, a set of preprocessing operations such as stemming or stop word elimination, and a set of weights for considered similarity measures. PSM computes the similarity between two concept names as the weighted average of values calculated by different selected similarity measures.

VMM enrich a concept name by extra information such as description field and neighbor names. Similarity between these enriched terms are then calculated using TF-IDF (Salton & McGill, 1986) technique.

RiMOM (Li, Tang, Li, & Luo, 2009) uses linear combination of modified Lin measure and a statistical measure. Falcon (Jian, Hu, Cheng, & Qu, 2005) employs a modified version of edit distance and combines the results by using the TF-IDF technique. CIDER (Gracia, Bernad, & Mena, 2011) uses Jaro Winkler measure to compute the similarity between concept names after enriching each concept name with its WordNet synonyms. CODI (Huber, Sztyler, & Noessner, 2011) combines various similarity measures such as Levenshtein and Jaro-Winkler through different methods. These methods include averaging, weighted averaging, maximizing. For weighted averaging methods, weights are calculated by a special learning method. AROMA (David, Guillet, & Briand, 2007) enhances the matching results by employing Jaro Winkler measure with a fixed threshold.

H-MATCH (Castano, Ferrara, & Montanelli, 2006) calculates the shortest path between the two entity names by using a thesaurus. LogoMap (Jimenez-Ruiz, Morant, & Grau, 2011) combines indexes, calculated by information retrieval technique, and anchor alignments to detect the matches among entities. LogoMap employs ISUB measure to compute anchor alignments confidences.

3. Approach

Within natural language, we use a vocabulary of atomic expressions and a grammar to construct well-formed and meaningful expressions as well as sentences. In the context of an ontology language the vocabulary is called *signature*, and can be defined as follows.

Definition 1 Signature. A signature $\mathbf{S}$ is a quadruple $\mathbf{S} = \langle \mathbf{C}, \mathbf{P}, \mathbf{R}, \mathbf{I} \rangle$ where $\mathbf{C}$ is a set of concept names, $\mathbf{P}$ is a set of object property names, $\mathbf{R}$ is a set of data property names, and $\mathbf{I}$ is a set of individual names. The union $\mathbf{P} \cup \mathbf{R}$ is referred to as the set of property names.

Definition 2 Similarity Search Algorithm σ. Given two ontologies O_1 and O_2 and their signatures $\mathbf{S}_1 = \langle \mathbf{C}_1, \mathbf{P}_1, \mathbf{R}_1, \mathbf{I}_1 \rangle$ and $\mathbf{S}_2 = \langle \mathbf{C}_2, \mathbf{P}_2, \mathbf{R}_2, \mathbf{I}_2 \rangle$, a similarity search algorithm σ is defined as:

$$\sigma(S, SimString) \rightarrow T \tag{15}$$

where $S = \mathbf{C}_2 \,|\, \mathbf{P}_2 \,|\, \mathbf{R}_2 \,|\, \mathbf{I}_2$ is the search space such that $T \in S$. $SimString \in \mathbf{S}_1$ is a search string. T type should be same as $SimString$, i.e. $SimString \in \mathbf{C}_1$ will lead to $T \in \mathbf{C}_2$ and so on. By reducing the problem with just considering one name from $\mathbf{S}_1$ as $SimString$, we tried to keep the algorithm more general, so it could be used by other applications such as search engines, which need to find a concept in an ontology similar to a search text. For the sake of the simplicity, in the followings, we only refer to concepts but similar methods could be applied to search in other parts of signatures.

In the following, we discuss the desired features for such similarity algorithm. First, the aim is to find the most specific concept in O_2 that is similar to $SimString$. This concept should not be more specific than $SimString$. Most of the semantic similarity measures that are defined based on edge counting are not applicable here since in such measures, concepts that have relation like sibling also receive a high similarity (due to close distance) and the direction of the relations does not matter. In fact, this requirement means that the algorithm should first try to find a concept very similar to $SimString$ and if failed, should try to find a concept similar to more general meaning of $SimString$. In other words, if $SimStringConcept$ be an ideational Concept in O_2 fully similar to $SimString$, then always $SimStringConcept \sqsubseteq foundConcept$. This feature is very important specially for aligning

two ontologies that expressed in different levels of granularity. Second, algorithm should not be sensitive to minor morphological differences such as suffixes and prefixes that have no effects on the meaning. In the same time, algorithm has to consider small none morphological changes that change the meaning totally (e.g., *Flood* and *Food*). Third, for some search applications such as search engines, high recall has a high importance, while for some other applications, like instance immigration between databases, more accurate results is definitely preferable. Consequently, the search algorithm has to be flexible and address different needs of recall and precision priority strategies. Fourth, the terminological search algorithm is the core component for most alignments systems, so having a high efficiency would directly improve the overall performance. Fifth, the algorithm should provide good level of stability because usually alignment algorithms are so sensitive to changes in their thresholds.

FINDSIMILAR(SimString,OntSearchList)

1: ▷ First tries to find resources that are similar to SimString directly
2: SimOntRes ← FINDLEXICALSIMILAR(SimString,OntSearchList ,IsubThrshld)
3: **if** SimOntRes = NIL **then**
4:    **if** SEMANTICFILTERACCEPTS(SimOntRes.LocalName,SimString) **then**
5:      **return** SimOntRes
6:    **end if**
7: **end if**
8: ▷ CreatingSearchMatrix
9: M ← WORDNETNUMBEROFMEANING(SimString)
10: SimMatrix ← BUILDEMPTYSIMILARITYMATRIX(M)
11: **for** i ← 0 to M − 1 **do**
12:   ADDTOROW(SimMatrix,i,WORDNETGETSYNONYMS(SimString,i))
13:   APPENDTOROW(SimMatrix,i,WORDNETGETHYPERNYMS(SimString,i))
14: **end for**
15: ▷ CalculateMostSimilars
16: CALCULATESIMILARITIES(OntSearchList , SearhMatrix)
17: CandidateArray ← BUILDARRAY(M)
18: **for** i ← 0 to M − 1 **do**
19:   CandidateArray [i] ← FINDCANDIDATE(SearhMatrix ,i)
20: **end for**
21: ▷ WordSenseDisambiguation
22: preferredMeaning ← WSD(SearchMatrix [i])
23: **if** CandidateArray [preferredMeaning][i] = NIL **then**
24:   **return** CandidateArray[preferredMeaning].MostSimilarOntRes
25: **end if**
26: ▷ If WSD failed
27: **for** i ← 0 to M − 1 **do**
28:   **if** CandidateArray [i] = NIL **then**
29:     **return** CandidateArray[i].MostSimilarOntRes
30:   **end if**
31: **end for**
32: ▷ Not found
33: **return** NIL

Figure 1. Similarity search algorithm

Proposed similarity search algorithm has been expressed in pseudo-code in Figure 1. It first tries to find the concepts in *OntSearchList* that are lexically very similar to *SimString*. This step is carried out by using a lexical search algorithm depicted in Figure 3. FindLexicalSimilar compares *SimString* to the all names exist in *OntSearchList*. Once it found the concept name that has the highest similarity to *SimString*, it compares their similarity value to a threshold, and if the similarity surpasses the needed threshold, it will return found similar concept. Otherwise, it will return null. For this direct comparisons, algorithm uses ISUB similarity measure

which, as reported (Stoilos et al., 2005), has a clear superiority in ontology name searching over other lexical measures. Once the lexical algorithm finished, the algorithm checks the returned concept (if exists any returned similar concept) by a semantic filtering algorithm (see Figure 4) to ensure that there is no semantic inconsistency between it and *SimString*.

As aforementioned in requirements, proposed algorithm should prevent the cases that two concepts are very similar lexically, but quite different semantically. We have added a SemanticFilteringAccepts method to fulfill this requirement. This method calculates the semantic similarity of two concepts by JIANG measure. If the similarity could not be determined, this method has not enough information to reject the similarity (showed in method by -1). If the proposed similar concepts have a semantic similarity lower than a threshold, Semantic Filtering Accepts will reject their similarity. Otherwise, the method accepts their similarity.

In the case that the algorithm fails to find a direct similar concept, it tries to find similar concepts by extending the *SimString* (lines 8-14). This algorithm extracts all *SimString* synonyms and hypernyms from WordNet and then creates matrix illustrated in Figure 2. Each row of this matrix represents one meaning of *SimString*. In fact, this algorithm not only tries to find concepts similar to *SimString*, but also tries to find concepts similar to synonyms and hypernyms of *SimString* as well. In each row of this matrix, all synonyms of the meaning that row represents, come first (from left) and all their hypernyms from most specific to most general come afterwards.
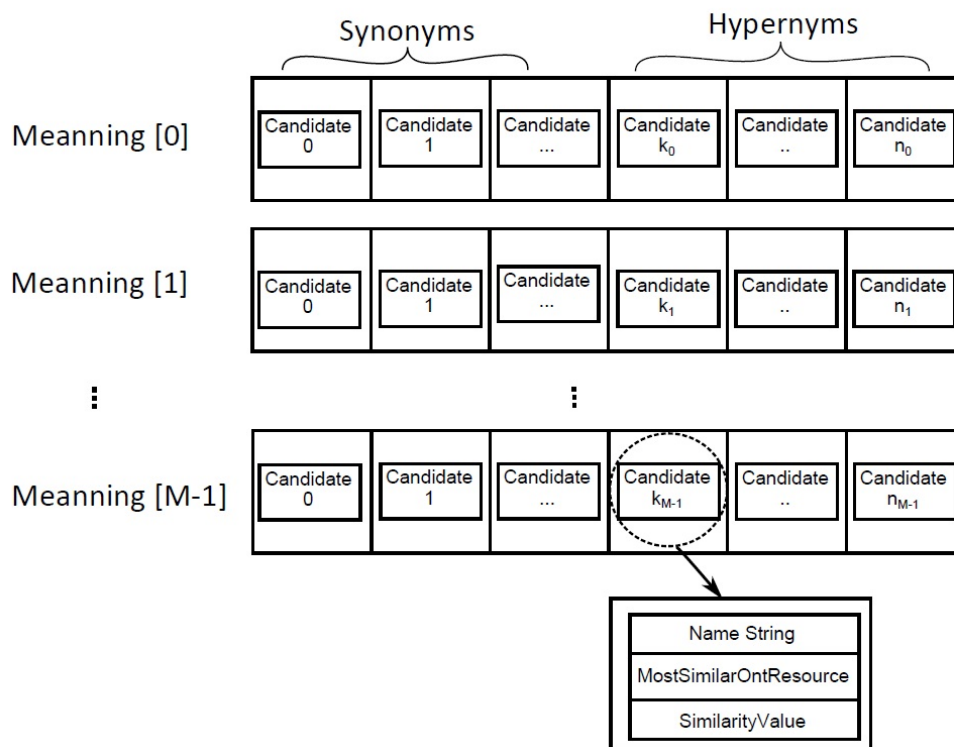

Figure 2. Searching matrix structure

Each *SimMatrix[i][j]* from this matrix can store one candidate for search result and contains three data fields; a *NameString* which algorithm tries to find most similar concept in *OntSearchList* to this and store it in *MostSimilarOntRes*, *MostSimilarOntRes* stores most similar concept to *NameString* which has been found so far, and finally, a *SimilarityValue* which shows the degree of similarity between *NameString* and *MostSimilarOntRes*.

CalculateSimilarites method (see Figure 5) compares each concept name from *OntSearchList* to all *NameString*s in matrix; if the concept name similarity to *NameString* is more than *SimilarityValue*, *MostSimilarOntRes* is replaced by this new concept, and *SimilarityValue* is also updated. Once all ontology concepts were compared to matrix candidates, each *SimMatrix[i][j]* would contain ontology concept that is most similar to containing *NameString*. For comparing the extracted synonyms and hypernyms to ontology concepts, we simply use the Levenshtein similarity measure since we know that ISUB has failed in previous step, and we want to use an alternative measure.

FINDLEXICALSIMILAR(SimString,OntSearchList,Thrshld)

 1: SimOntRes ← NIL
 2: SimOntResSimValue ← 0.0
 3: **for all** OntRes ∈ OntSearchList **do**
 4:    Name ← OntRes · LocalName
 5:    Sim ← ISUBSIMILARITY(SimString,Name)
 6:    **if** Sim > SimOntResSimValue **then**
 7:      SimOntRes ← OntRes
 8:      OntResSimValue ← Sim
 9:    **end if**
10: **end for**
11: **if** Sim > Thrshld **then**
12:    **return** SimOntRes
13: **end if**
14: **return** Nil

Figure 3. Find lexical similar algorithm

As noted earlier, each row of matrix includes all synonyms and hypernyms of one meaning of *SimString*, in order which most specific names at left to most general at right. By considering this order, a *DistanceThreshold*, values between [0, 1], specifies the ratio of each row from left that algorithm is permitted to scan for finding similar concepts. In other words, lower *Distance Threshold* means to consider less general hypernyms. Conversely, higher *Distance Threshold* means to consider more general hypernyms. Using this parameter can provide more flexibility which specified in requirements.

Following this, algorithm chooses at most one candidate from each row of this matrix by calling FindCandidate (see Figure 5) and puts them in *CandidateArray*. In the Find Candidate one row of matrix, which represents one meaning of *SimString,* is scanned from left to right to find the first candidate that has a Similarity Value higher than SimThreshold. If found candidate is rejected by Semantic Filter Accepts method, the selected candidate is simply ignored and scanning in the same row will be continued to find another possible candidate.

SEMANTICFILTERINGACCEPTS(OntResName,SimString)

 1: Sim ← JIANGSIMILARITY(OntResName,SimString)
 2: **if** Sim = −1 or Sim > SemanticFilteringLowThreshold **then**
 3:    **return** true
 4: **else**
 5:    **return** false
 6: **end if**

Figure 4. Semantic filtering accepts algorithm

Once the algorithm finished establishing the possible candidate list, it can return the candidate list to let the application to choose the suitable candidate upon its preferred context. Alternatively, it can do the word sense disambiguation (WSD) (Navigli, 2009) by itself and then return the selected candidate. In this algorithm we have implemented a straightforward WSD technique using context knowledge which has been accumulated in *SearchMatrix*. Each row of *SearchMatrix* contains synonyms and hypernyms of one *SimString*'s meaning. As aforementioned, after calling the CalculateSimilarites method, each element of *SearchMatrix* is populated by the most similar concept in addition to the similarity value. The WSD method uses average of similarity values in each row as measure of its relatedness to the ontology. The higher row's similarity values average, the higher probability that the row represents the wanted meaning of *SimString* in ontology. The rationale behind this rule is that having higher average shows that the row has more commonality with the ontology concept names.

CALCULATESIMILARITIES(OntSearchList,SearchMatrix)

1: **for all** OntRes ∈ OntSearchList   **do**
2:    **for** i ← 0 to M − 1 **do**
3:      **for** j ← 0 to ROWSIZE(SimMatrix, i) × DistanceThreshold   **do**
4:        Sim ← LEVENSIMILARITY(Name,SimMatrix[i][j] · NameString)
5:        **if** Sim > SimMatrix[i][j] · SimilarityValue   **then**
6:          SimMatrix[i][j] · SimilarityValue ← Sim
7:          SimMatrix[i][j] · MostSimilarOntRes ← OntRes
8:        **end if**
9:      **end for**
10:    **end for**
11: **end for**

Figure 5. Calculate similarities algorithm

Finally, the algorithm will return the selected candidates, and if the WSD fails to choose any candidate (i.e. the selected row does not contain any candidate), the algorithm tries to select the first candidate based on their order in the WordNet because the synsets in the WordNet are sorted by their usages' frequencies (Fellbaum, 1998).

FINDCANDIDATE(SearchMatrix, Row, Thrshld, SimString)

1: **for** j ← 0 to ROWSIZE(SearchMatrix, i) × DistanceThreshold   **do**
2:    **if** SearchMatrix[Row][j] · SimilarityValue > LevThrshld **then**
3:      OntResName ← SearchMatrix[Row][j].MostSimilarOntRes.LocalName
4:      **if** SEMANTICFILTERINGACCEPTS(OntResName ,SimString) **then**
5:        **return** SearchMatrix[Row][j]
6:      **end if**
7:    **end if**
8: **end for**
9: **return** Nil

Figure 6. Find candidate algorithm

4. Experiments

For examining performance of implemented search function, OAEI benchmark (Euzenat et al., 2011) 101 and 205 datasets have been used to compare suggested search algorithm to different search algorithms based on variety of syntactical and semantical similarity measures. OAEI benchmark 205 dataset has been designed to show effectiveness of ontology matching algorithms in using string similarity. In our proposed algorithm, three different syntactical and semantical measures have been used: the ISUB measure as the main syntactical measure, the Levenshtein as the auxiliary syntactical measure, and the JIANG as the semantical measure for filtering mechanism. In our test scenarios we compared results of our approach to use of these three measures separately. In addition, to investigate performance of our algorithm more comprehensively, some other similarity measures from different group of semantic measures have been exploited in our test scenarios. These similarity measures include Lin, Resnik, and Pirro. It has to be noted that we cannot compare results of our algorithm, which just use name similarity, with full ontology matching algorithms that use many various matching algorithm.

Furthermore, we carried out two extra tests that consider aggregation of these measures as well. In the first aggregation scenario, the average of all these measures is been calculated while, in the second scenario, these three measures are prioritized according to their order in our algorithm. We also added some more experiences in order to investigate some other aspects of proposed algorithm such as its stability and effectiveness of features like semantic filtering or word sense disambiguation.
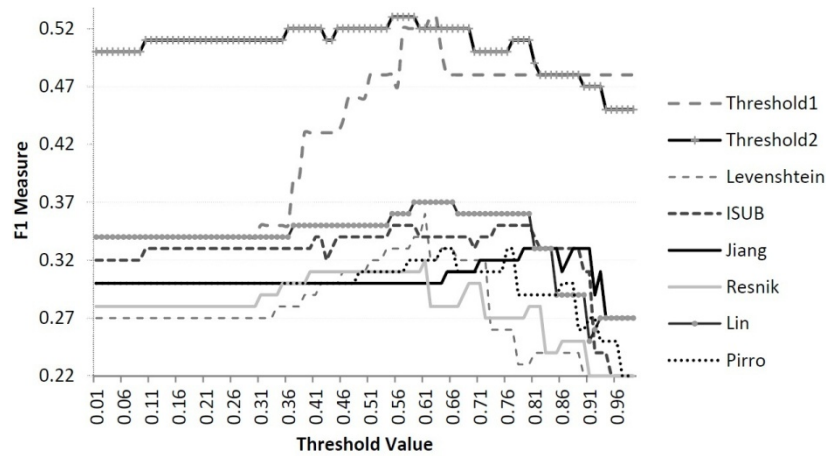
Figure 7. F1 measures of achieved results have been compared to other search algorithms based on well-known semantic and string similarity measures

Description: Threshold1 shows the results which has been achieved by changing the first threshold of algorithm, and Threshold 2 shows the results for changing second threshold of algorithm.
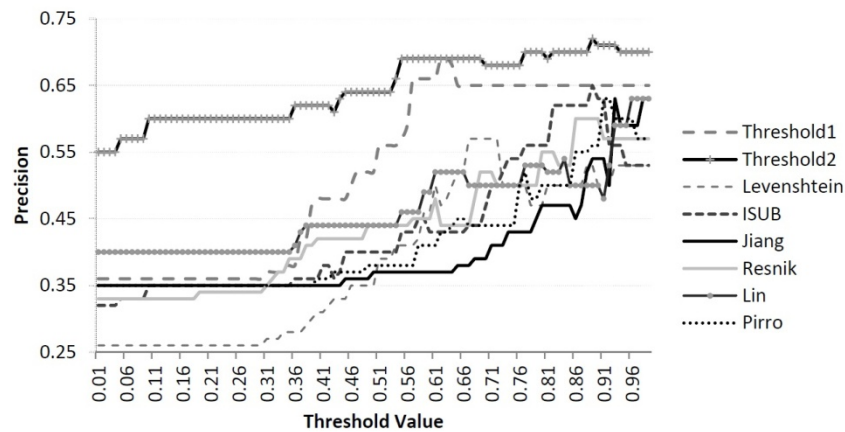

Figure 8. Precisions of achieved results have been compared to other search algorithms based on other well-known semantic and string similarity measures

The proposed algorithm, like most of the other matching algorithms, needs some setting parameters. First, it needs the threshold used for the main lexical similarity search algorithm (see Figure 3) which uses the ISUB similarity measure, so we will refer to it as ISUB threshold or sometimes simply threshold1. Second, it requires the threshold for semantic filtering (see Figure 4) which uses the JIANG similarity measure. Third, it exploits a threshold for finding candidates for each row (see Figure 6). Since this algorithm uses Levenshtein measure, we refer to this threshold as Levenshtein threshold. Finally, there is also a distance threshold which put a limitation on the percentage of hypernyms which considered in each row of the *Search Matrix*. We focus more on the first and second thresholds, but we also accomplish some experiences on the effects of changing other thresholds on the algorithm performance.

In Figure 7, Figure 8, and Figure 9, performance of proposed algorithm has been compared with the matchers have made based on other measures in the term of F1 measure, precision, and recall respectively. For each of these measures, we have developed a basic matcher and applied it to the data set. In each matcher, regarding threshold was set from 0.0 to 0.99 with the step of 0.01. For our algorithm, the *Thrshold 1* shows the results of changing ISUB threshold, and the *Thrshold 2* shows the effects of changing in Levenshtein threshold while keeping the other thresholds in the fixed best values. These results show that our algorithm outperforms all other measures when they are used separately in the term of result quality factors. These results also reveal that the algorithm is more sensitive to changes in the first threshold other than changes in the second threshold.
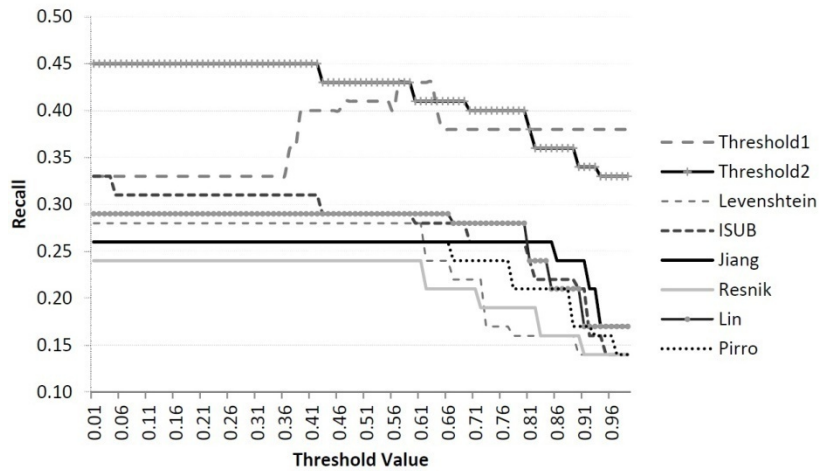
Figure 9. Recall results of proposed method have been compared to other search algorithms based on other well-known semantic and string similarity measures

The previous experiment has been repeated to compare the run time needed for each developed matcher. Although the developed algorithm running time is not comparable with syntactical measures such as Levenshtein (run time was less than 20 MS and has not been shown), the Figure 10 shows that it has a running time far better than other semantic measures.
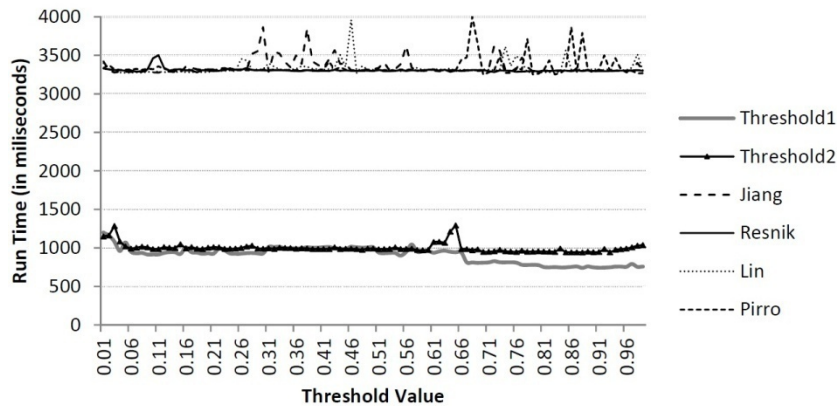

Figure 10. Run time comparison of proposed algorithm to other search algorithms based on well-known semantic measures

The other concern about the using of an algorithm with numbers of thresholds is the stability of the algorithm. In other words, algorithm has to show that its performance would not change sharply by changing its thresholds. Figure 11 shows the changes in F1 measure of the results due to changes in both first and second thresholds. These results demonstrate that algorithm is not so sensitive to its thresholds changes. It shows a good level of stability because the majority area of the curve has a high F1 measure, and the curve is smooth without sharp drops or raises.

One feature of the proposed algorithm is the using of a *distance threshold* to put a limit on the percentage of used hypernyms in each line of search matrix (see Section 3). Figure 12 illustrates the effects of changing *distance threshold* from 0.0 to 1.0 while all other parameters such as similarity and filtering thresholds are in their best values. This results show that after a fluctuation, F1 measure has been increased to its highest value when the *distance threshold* is 0.81 and then has been decreased again afterward. This experience shows that using more hypernyms will improve the results, but employing very general hypernyms shows fewer impacts on the results.
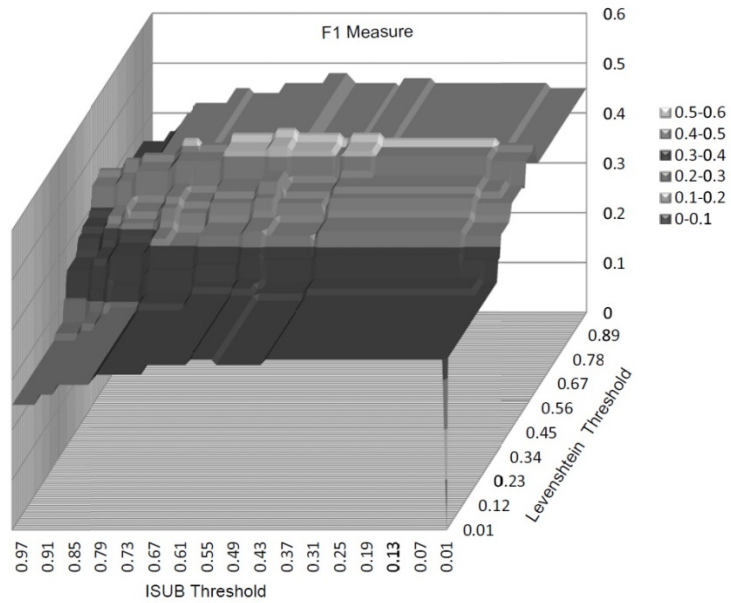
Figure 11. Searching for best combinations of both thresholds of algorithm

Description: ISUB axis represents the values for the first threshold of algorithm and the Levenshtein axis shows the values for the second threshold of algorithm.
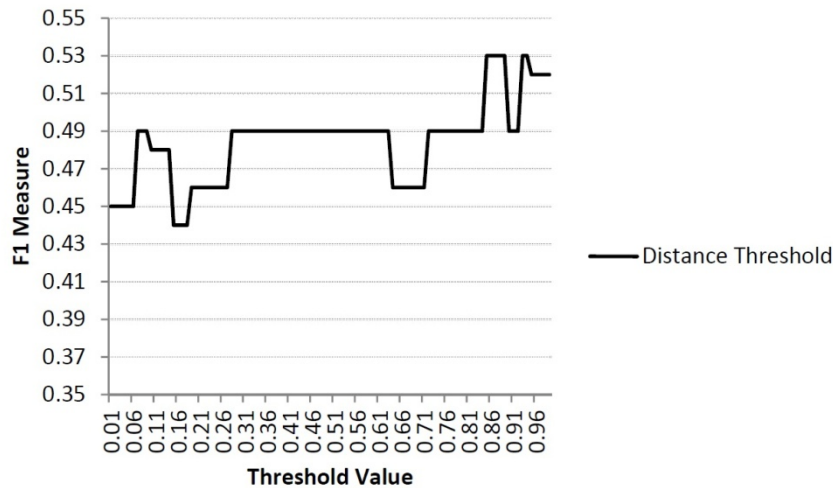

Figure 12. The effect of changing distance threshold parameter on the results of algorithm in the term of F1 measure

As discussed earlier (see Section 3), this algorithm applies a filtering mechanism to eliminate the false similar causes that the two words are lexically very similar but semantically different. Figure 13 shows the significant improvement which has been achieved by employing this mechanism independent from the used similarity thresholds. In this experiment, we changed the first ISUB threshold from 0.0 to 1.0. However, the runtime values illustrated in the Figure 14 show that this mechanism has made the algorithm almost three times slower.

Figure 15 demonstrates that applying the simple proposed word sense disambiguation mechanism can improve the results. Additionally, as illustrated in Figure 14, the word sense disambiguation due to its simplicity will not put significant overload on the main algorithm.
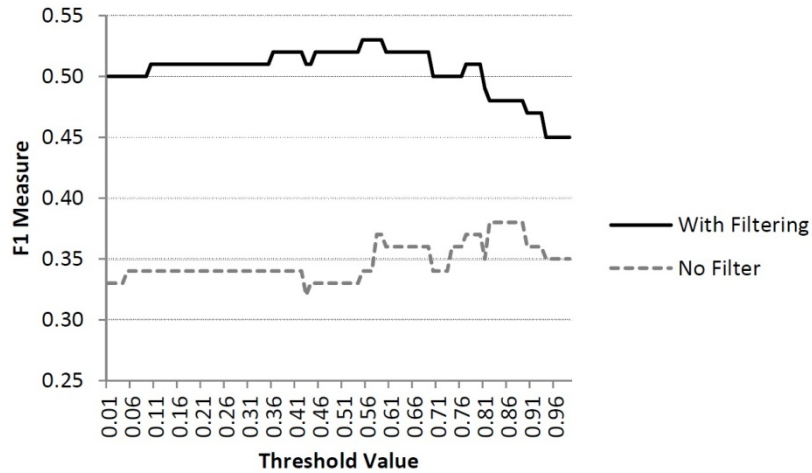
Figure 13. The achieved results of algorithm which uses filtering mechanism compared to results achieved without using filtering

Figure 14. The cost of using filtering mechanism has been showed by comparing run time of algorithm that uses filtering mechanism and the run time of algorithm that does not use filtering

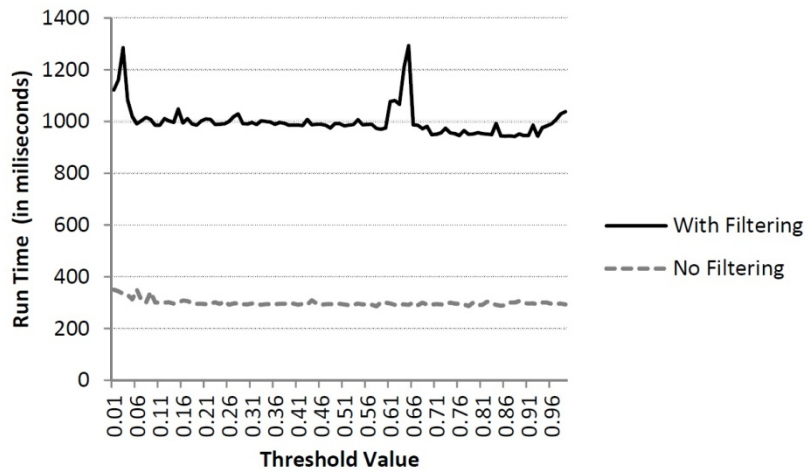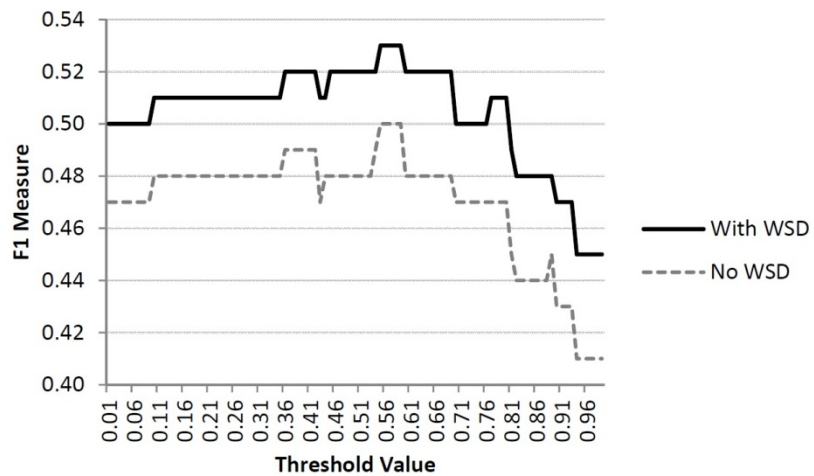

Figure 15. Results of algorithm which use word sense disambiguation mechanism compared to results has been achieved without using word sense disambiguation
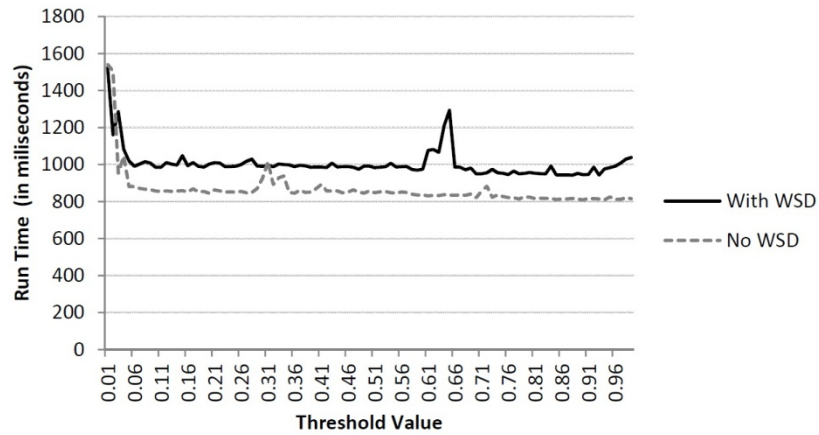
Figure 16. The cost of using WSD mechanism has been showed by comparing run time of algorithm that uses WSD mechanism with run time of the algorithm does not use WSD
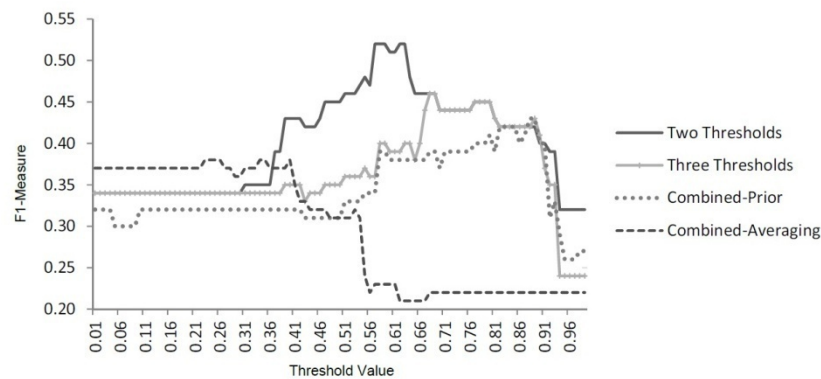

Figure 17. The results of proposed algorithm compared to other algorithms that combine different measures

Our algorithm uses three measures which are Isub, Levenshtein, and JIANG. In the last experience, it has been tried to compare suggested algorithm (*Three Threshold* in Figure 17) to some other previous methods that also combine those measures to improve the overall performance. Two conventional approaches are: comparing average of all three different measures (*Combined-Averaging* in Figure 17) and using different measure with priority (*Combined-Prior* in Figure 17). In using with priority case, it begins to calculate the similarity with the first measure and then compares it to the threshold. If the similarity value is less than the threshold, it will use the next measure. In the other hand, this method considers that two words are similar if at least one calculated similarity is higher than the threshold. It should be noted that the aforementioned priority just has influence on the algorithm runtime and not the results. Setting some thresholds to the best values in previous experiments can be subject of argumentation, so it has been avoided to use any specific values for the needed thresholds in this last experiment. In the configuration of our algorithm's thresholds, instead of using fixed value for each threshold, we have just used a same variable value for all needed thresholds.

Figure 17 compares the results for those three different approaches (*Two Threshold* results are the results achieved by a fixed value for semantic filtering mechanism and could be used as an estimation of proposed algorithm final results). This figure shows that in the first, the average approach demonstrates a better performance, but with increase in the threshold this method performance has been dropped to the lowest values such that it is not comparable to other methods. Priority based approach has also demonstrated lower F1 measure in comparison with proposed algorithm. This experiment shows that even without taking in consideration any threshold setting the proposed algorithm has surpassed its competing approaches.

5. Conclusions

In this paper, we proposed a novel terminological search algorithm which tries to find a concept Name (property or individual) similar to an input Search String in a given ontology. This search algorithm is a basic building

block for many semantic applications such as ontology matching systems and ontology search engines. As we showed in related works, while there exist a lot of ontology matching approaches and also many proposed similarity measures, little attentions have been paid to develop similarity search algorithms that exploit different similarity measures. Such algorithm can use various similarity measures to sum up their advantages and reduce effects of their weakness.

Our suggested algorithm extends the input search string by creating a matrix from its synonym and hypernyms which have been extracted recursively from WordNet such that each row of this matrix represents one meaning. Each row includes synonym from left and then most specific hypernyms and finally more general hypernyms come afterwards. The algorithm first uses the ISub measure to find similar concepts, and if failed to find any lexically similar concept, it will continue to search the extension matrix. In both cases, algorithm uses JIANG semantic similarity measure to detect wrong candidates which are lexically similar and semantically different. For coping with the word polysemy problem, algorithm use a simple word sense disambiguation method based on average relatedness of each row of the search matrix. The algorithm employs and combines different kind of similarity measures in different situations to achieve a higher performance, accuracy, and stability compared to previous methods which either use one similarity measure or combine them in a naive ways such as averaging.

For algorithm evaluation, we used OAEI Bench mark data set and the achieved results showed the superiority of proposed algorithm and effectiveness of different suggested mechanism such as word sense disambiguation and semantic filtering mechanism.

There are some potential limitations in this study. First, the proposed algorithm is not well suited for search names that have many parts since it does not employ any tokenization method. This is an important requirement which should be implemented before embed this algorithm in a real ontology matching system because some major ontologies especially from medical world usually use long concept names. Nonetheless, in some applications such as semantic search engines this search algorithm in its current implementation could be very useful. It should be noted that in complex matching task (Ritze, Volker, Meilicke, & Sváb-Zamazal, 2010) usually finding concepts similar to a part of a compound name is an important basic operation.

Second, this algorithm mainly relies on the WordNet to enrich the search string while ontologies are defined in different domains and need taxonomy and background knowledge that better fit their requirements. This algorithm also could exploit the content of each concept in implementing task such as WSD or semantic filtering.

In future works, we will broaden our approach to use more linguistics techniques such as tokenization, stop word reduction, and stemming to make it more suitable to cope with ontologies that have long compound names. We would also like to generalize algorithm such that employing knowledge sources other than WordNet become possible in different situations. Specially, employing ontologies themselves in building extension matrix could be an interesting improvement. Another future research direction is to exploit more sophisticated and state of the art word sense disambiguation techniques.

References

Berners-Lee, T., Hendler, T., & Lassila, J. (2001). The Semantic Web. *Scientific American*. http://dx.doi.org/10.1038/scientificamerican0501-34

Castano, S., Ferrara, A., & Montanelli, S. (2006). Matching Ontologies in Open Networked Systems: Techniques and Applications. *J. Data Semantics V.,* 25-63.

Cross, V., & Hu, X. (2011). Using Semantic Similarity in Ontology Alignment.

Cruz, I. F., Antonelli, F. P., & Stroe, C. (2008). *Efficient Selection of Mappings and Automatic Quality-driven Combination of Matching Methods*. Paper presented at the OM.

David, J., Guillet, F., & Briand, H. (2007). Association Rule Ontology Matching Approach. *International Journal of Semantic Web Information Systems, 3*(2), 27-49. http://dx.doi.org/10.4018/jswis.2007040102

Department, Z. W., & Wu, Z. (1994). Verb Semantics And Lexical Selection. *In Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, 133-138.

Euzenat, J., Meilicke, C., Stuckenschmidt, H., Shvaiko, P., & dos Santos, C. T. (2011). Ontology Alignment Evaluation Initiative: Six Years of Experience. *J. Data Semantics, 15*, 158-192.

Euzenat, J., & Shvaiko, P. (2007). *Ontology matching*. Springer-Verlag.

Euzenat, J., & Valtchev, P. (2003). *An integrative proximity measure for ontology alignment*. Paper presented at

the The 1st Intl. Workshop on Semantic Integration.

Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*.

Gracia, J., Bernad, J., & Mena, E. (2011). *Ontology Matching with CIDER:evaluation report for OAEI 2011*. Paper presented at the The 6th International Workshop on Ontology Matching.

Huber, J., Sztyler, T., & Noessner, J. (2011). *CODI: Combinatorial Optimization for Data Integration*. Paper presented at the The 6th International Workshop on Ontology Matching.

Jaro, M. A. (1995). Probabilistic Linkage of Large Public Health Data Files. *Statistics in Medicine, 14*, 491-498. http://dx.doi.org/10.1002/sim.4780140510

Jean-Mary, Y. R., Shironoshita, E. P., & Kabuka, M. R. (2009). Ontology matching with semantic verification. *J. Web Sem., 7*(3), 235-251. http://dx.doi.org/10.1016/j.websem.2009.04.001

Jian, N., Hu, W., Cheng, G., & Qu, Y. (2005). *FalconAO: Aligning Ontologies with Falcon*. Paper presented at the Integrating Ontologies.

Jiang, J., & Conrath, D. (1997). *Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy*. Paper presented at the The International Conference Research on Computational Linguistics (ROCLING).

Jimenez-Ruiz, E., Morant, A., & Grau, B. C. (2011). LogMap results for OAEI 2011.

Leacock, C., & Chodorow, M. (1998). *Combining local context and WordNet similarity for word sense identification*. Paper presented at the MIT Press.

Levenshtein, V. I. (1966). *Binary codes capable of correcting deletions, insertions, and reversals*. Paper presented at the Soviet Physics Doklady.

Li, J., Tang, J., Li, Y., & Luo, Q. (2009). RiMOM: A Dynamic Multistrategy Ontology Alignment Framework. *IEEE Transactions on Knowledge and Data Engineering, 21*, 1218-1232.

Lin, D. (1998). *An Information-Theoretic Definition of Similarity*. Paper presented at the ICML.

Navigli, R. (2009). Word sense disambiguation: a survey. *ACM COMPUTING SURVEYS, 41*(2), 1-69. http://dx.doi.org/10.1145/1459352.1459355

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology, 48*(3), 443-453. http://dx.doi.org/10.1016/0022-2836(70)90057-4

Pirro, G. (2009). A semantic similarity metric combining features and intrinsic information content. *Data Knowl. Eng., 68*, 1289-1308. http://dx.doi.org/10.1016/j.datak.2009.06.008

Rada, R., Mili, H., Bicknell, E., & Blettner, M. (1989). Development and application of a metric on semantic nets. *IEEE Transactions on Systems, Man, and Cybernetics, 19*(1), 17-30. http://dx.doi.org/10.1109/21.24528

Resnik, P. (1995). *Using Information Content to Evaluate Semantic Similarity in a Taxonomy*. Paper presented at the In Proceedings of the 14th International Joint Conference on Artificial Intelligence.

Ritze, D., Volker, J., Meilicke, C., & Sváb-Zamazal, O. (2010). *Linguistic Analysis for Complex Ontology Matching*. Paper presented at the Proceedings of the 5th International Workshop on Ontology Matching (OM-2010), Shanghai, China, November 7, 2010.

Romero, M. M., Vázquez-Naya, J. M., Loureiro, J. P., & Ezquerra, N. (2009). Ontology Alignment Techniques *Encyclopedia of Artificial Intelligence* (pp. 1290-1295): IGI Global.

Salton, G., & McGill, M. J. (1986). *Introduction to Modern Information Retrieval*: McGraw-Hill, Inc.

Stoilos, G., Stamou, G., & Kollias, S. (2005). A String Metric for Ontology Alignment. *The Semantic Web – ISWC 2005, 3729*, 624-637.

Tversky, A. (1977). *Features of Similarity*. Paper presented at the Psychological Review.

Winkler, W. E. (1999). *The state of record linkage and current research problems*.

Analysis and Modelling of Effects of Traffic Light Operations Variability to Violation Rates at Junction

Fabio Galatioto[1], Tullio Giuffrè[2], Margaret C. Bell[1] & Giovanni Tesoriere[2]

[1] School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, United Kingdom

[2] Faculty of Engineering and Architecture, Kore University of Enna, Enna, Italy

Correspondence: Tullio Giuffrè, Faculty of Engineering and Architecture, Kore University of Enna, Cittadella Universitaria, Enna 94100, Italy. Tel: 39-935-536-356. E-mail: tullio.giuffre@unikore.it

Abstract

Previous research, based on potential conflicts analysis, has provided a quantitative evaluation of 'proneness' to red-light running (RLR) behaviour at urban signalised intersections (Giuffrè & Rinelli, 2006) varying geometric, traffic flow and driver characteristics. Recent study (Bell, Galatioto, Giuffrè, & Tesoriere, 2012) demonstrated the potential to use micro-simulation models to evaluate the 'proneness' to RLR behaviour at urban signalised intersections in Italian cities, varying flow characteristics and stop line distances.

The micro-simulation, although at its early stages of development, has shown promise in its ability to model unintentional RLR behaviour and to evaluate alternative junction designs on proneness.

In order to make more robust the new modelling framework, the need to demonstrate the transferability of the modelling approach has been addressed in this paper by using a four arms junction in Enna (Italy), where an extensive period of continuous video recording has been carried out. Moreover, in collaboration with the local Police, different cycle and green time settings have been implemented in order to measure the effects on the RLR rates. Then the measured RLR rates have been evaluated and compared to the modelling results as a validation exercise. In this way the prediction capability of the proposed potential conflict model has been extended and improved.

Keywords: urban intersection, traffic light operation, violation analysis, microsimulation, road design evaluation, proneness to RLR

1. Introduction

Evidence of the human and financial cost to individuals and Governments (Retting & Williams, 1996) of Red-Light Running (RLR) at signalised junctions has led to several studies over the recent decades. Also, various research (Mohamedshah, Chen, & Council, 2000; Milazzo, Hummer, & Prothe, 2001) analysed the effects of the RLR phenomenon on road safety. They all highlight both a very high frequency and the respective magnitude of the consequences (Wissinger, 2000).

Recently, the European Parliament and Europe Council (CEC, 2008) identified red-light running as one of the greatest causes of crashes, along with speeding, drunk-driving, and not using a seatbelt. This is confirmed also in other areas around the world such as U.S. where a study of accident data from four States (Mohamedshah et al., 2000) shown that RLR crashes account for up to 20 percent of the total crashes at urban signalised intersection.

As referring by Giuffrè and Rinelli (2006) and according to Elmitiny, Yan, Radwan, Russo and Nashara (2010) proneness to RLR behaviour can be associated to two main factors which determine - when circumstances make it possible - the choice to stop, or to proceed, at the junction. These are:

-*human factors*, related to population characteristics;

-*road factors*, depending on traffic, geometric and operational characteristics at the intersection.

Several behavioural studies have highlighted the role of the 2 main contributing features on RLR in terms of "dilemma zone" (Allos & Al Haidithi, 1992) and situations that make RLR possible (Bonneson, Brewer, & Zimmerman, 2001). Data analysis carried out by Campbell, Smith and Najm (2004) allowed building some logistic regression models to examine violators' vehicle speed and time after red when they crossed the

intersection after the onset of the red signal.

Particularly in urban areas, many studies have been developed both on safety at signalised intersections - refer to Giuffrè, Granà, Marino, and Corriere (2011) - and on the set up of the traffic light signal timings, that is crucial to ensure that driver are likely to respect the signals. The FHWA (2003) suggests that the traffic light signal timings should be reviewed and updated regularly (every two years) in order to adapt the timings to the modified traffic demands. In particular, well-defined traffic light signal timing can reduce driver frustration resulting from extremely short or extremely long cycle duration (Kennedy & Sexton, 2009).

The setup of the traffic light signal timings and phases is based on the specific characteristics of the junction and the single approaches. There are philosophies and studies that indicate both shorter and longer cycle duration for reducing violations at signalised junction (Sze, Wong, Pei, Choi, & Lo, 2011; Koh & Wong, 2007). A driver that expects to wait for a not excessive period of time may be less prone to accelerate to beat the amber or to run during the red phase. Different studies, summarised by (Chodur, Ostrowski, & Tracz, 2011) shown the effects of cycle length vary on traffic and driver.

Although acknowledging that the human factors component of RLR is important, different studies in the past have considered the qualitative influence of both human, vehicle and road factors. From an engineering point of view the availability of qualitative data alone is insufficient in order to establish appropriate countermeasures against the RLR phenomenon on the decision making process. For this reason there is also a strong need to evaluate quantitatively the proneness to RLR. However, availability of this quantitative data is scarce.

Recent application from the authors Bell, Galatioto, Giuffrè, and Tesoriere (2010), and Bell et al. (2012) have tried to extend the traditional empirical and probabilistic approaches by assessing the proneness to RLR using a micro-simulation model in a signalised junction. In this first attempt it was demonstrated how the modelling tool has potential capability to evaluate the effect of geometric designs, operational conditions and traffic flow volumes on the unintentional RLR violation rates. Because of the potential value for traffic operators and Local Authorities of the extended use of existing micro-simulation models, already setup in their areas of competences and cities, to identify potential high risk junctions and road layouts, this paper demonstrates how the same above mentioned approach can be transferred into a different study site.

Moreover, acknowledging that RLRs behaviour is influenced by the traffic light timings, usually setup to optimise traffic movements at junctions, and related delays and stops, the authors in collaboration with the local police defined two sets of traffic light cycles and several green times changes in order to measure how the RLR rates will be affected.

Finally, this paper seeks to calibrate and validate a micro-simulation tool that could be used to minimise proneness of RLR at the design stage. The junction used for the research is described in the following case study section. The paper will continue describing the potential for the validated micro-simulation model to predict the proneness of links / streams to RLR. Finally, the micro-simulation method validated here will provide evidence of its potential use to explore, at a design stage, different geometric layouts and traffic light control strategies.

2. Case Study

A road intersection (see Figure 1) in Enna, Italy, was studied in detail over an extensive period of 13 days.

Numerous business activities characterise the study area, such as bars, offices overlooking arms 2 and 3, while in the left side on the corner between arm 1 and 4 there is the elementary school "De Amicis". Other businesses (such as banks, wine bars, restaurants) are in close proximity of the intersection, increasing the pedestrian movements especially during day time. Moreover, the intersection connects the city to the area used for the weekly market placed at the Europa square.

The location and type of businesses in the area studied are identified in detail in Figure 1 (right).

The red-light running information have been collected by manual on-street measurements and validated using video surveys used to monitor traffic flows and speed over the survey period.

The junction is a typical four arm intersection with a main road, namely Viale Diaz, indicated in Figure 1 with arm 1 and 3, both two ways and one lane road; and a secondary road, namely Via Libertà, with arm 2 and 4, where arm 2 is a two lanes one way road and arm 4 a two ways road with one lane for each direction.

The intersection as represented in Figure 1 (left) is an isolated signalised junction. Only other priority junctions are present in close proximity of the junction studied, for clarity in Figure 1 (left) the 2 closer priority junctions in each arm have been represented. This condition provides an arrival pattern which is not dependent on other traffic light timings avoiding platoon arrival patterns as well.

Figure 1. Type of junctions (left) and layout of the signalised intersection studied in Enna, with characterisation of business activities (right)

The duration of the traffic-light cycle (C) before the modification made for the purpose of this study, was equal to 84 seconds. During the monitoring period a variation of the traffic-light cycle has been setup according to the following modifications:

- increase of traffic light cycle from 84 to 94 seconds;
- variation of green and red time keeping the duration of the cycle constant.

In Table 1 are presented for the two directions the different traffic light timings setup through the monitoring period.

Table 1. Traffic signal phasing during the observation period

day	Main direction - Viale Diaz					Secondary direction - via Libertà				
	green	red	AR	yellow	Cycle	green	red	AR	yellow	Cycle
1	30	49	3	5	84	40	39	3	5	84
2	55	34	3	5	94	25	64	3	5	94
3	45	44	3	5	94	35	54	3	5	94
4	55	34	3	5	94	25	64	3	5	94
5	50	39	3	5	94	30	59	3	5	94
6	50	29	3	5	84	20	59	3	5	84
7	25	64	3	5	94	55	34	3	5	94
8	40	39	3	5	84	30	49	3	5	84
9	40	39	3	5	84	30	49	3	5	84
10	40	39	3	5	84	30	49	3	5	84
11	40	39	3	5	84	30	49	3	5	84
12	40	39	3	5	84	30	49	3	5	84
13	40	39	3	5	84	30	49	3	5	84

The traffic light cycle consists of 2 phases along the two roads. During the monitoring period the traffic light cycle was modified for a period of 7 days while for the remaining 6 days was maintained identical to the original standard duration.

The traffic light duration has been modified according to Table 1 timings mainly for the purpose of testing the variability of violations and unintentional RLR rates as a consequence of the different traffic light settings. As

the volumes of traffic for the different arms were on the range 100 up to 1,000 vehicles per hour, with the highest flow coming from Via Liberta' arm 2, which has 2 lanes, the capacity of each link was considerably greater than the maximum flow measured with expected Level of Services A or B. However, in this paper was not assessed the performance of each traffic light setting in term of Level of Services.

With reference to Figure 2 there is a clear underlying trend between day 1 and day 8 of progressive reduction of the number of violations in the main direction, while in the remaining period, when the traffic light setup was maintained constant and equal to day 1, it shows a progressive increment of the violation toward a rate similar to the beginning of the survey period.

In Figure 2 are presented the number of violations and the cycle times (in seconds) along the main and secondary roads.
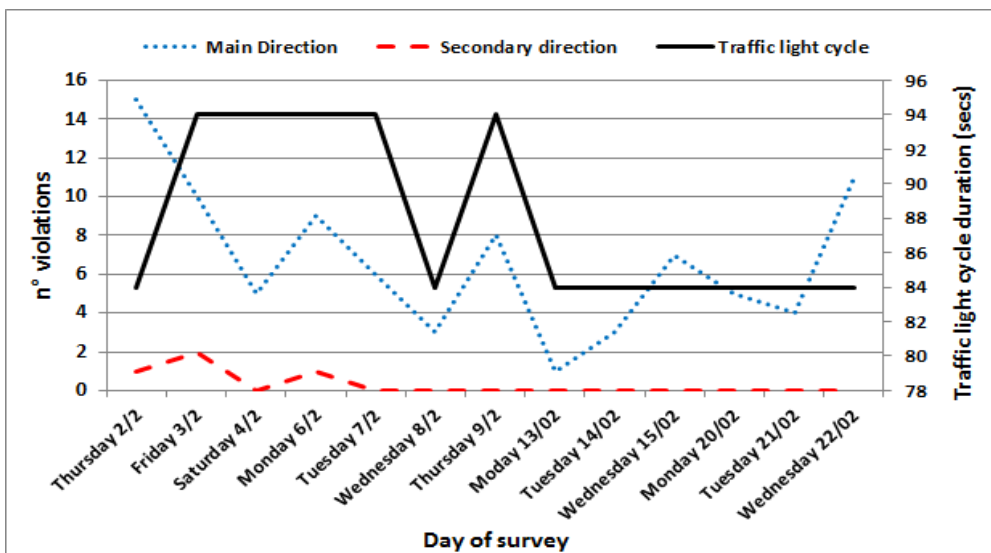

Figure 2. Number of violations and related traffic light cycle during the 13 days of observation along the main and secondary roads

During weekdays the daily traffic of the junction is 11,500 vehicles, while on the main road is 7,300 veh/day and 4,200 veh/day on the secondary road.

3. On-site Measurements

After the validation exercise, it was possible to assume that the actual RLRs events were equal to the violations observed. Table 2 aggregates the violation rates based on violation/hour, violation/1,000 vehicles and violation per 10 000 vehicles-cycle for the selected intersection.

The collection of the data, using video recording, postprocessing and manual techniques was carried out on a daily basis for 11 hours of relief on both directions from 7:30 to 13:30 and from 15:30 to 20:30.

Table 2. Violation rates for the selected intersection over the observation period

		Main direction				Secondary direction				
	day	N. violation (11 hours)	Viol / hour	Viol / 1,000 veh	Violations per 10 000 veh-cycle	N. violation (11 hours)	Viol / hour	Viol / 1,000 veh	Violations per 10 000 veh-cycle	Traffic Light Cycle
1	*Thu 02/02/2012*	15	1.364	1.763	0.037	1	0.091	0.341	0.007	84
2	*Fri 03/02/12*	10	0.909	1.154	0.027	2	0.182	1.134	0.027	94
3	*Sat 04/02/12*	5	0.455	0.747	0.018	0	0	0	0	94
4	*Mon 06/02/12*	9	0.818	0.938	0.022	1	0.091	0.445	0.011	94
5	*Tue 07/02/12*	6	0.545	0.753	0.018	0	0	0	0	94

6	*Wed 08/02/12*	3	0.273	0.433	0.009	0	0	0	0	84
7	*Thu 09/02/12*	8	0.727	0.928	0.022	0	0	0	0	94
8	*Mon 13/02/12*	1	0.091	0.157	0.003	0	0	0	0	84
9	*Tue 14/02/12*	3	0.273	0.402	0.009	0	0	0	0	84
10	*Wed 15/02/12*	7	0.636	0.954	0.020	0	0	0	0	84
11	*Mon 20/02/12*	5	0.455	0.624	0.013	0	0	0	0	84
12	*Tue21/02/12*	4	0.364	0.438	0.009	0	0	0	0	84
13	*Wed 22/02/12*	11	1.00	1.195	0.025	0	0	0	0	84

4. Microsimulation Tool Description

Microscosimulation models are considered the most powerful and appropriate traffic analysis tools when the aim is to reproduce movements of individual vehicle in a network (through junctions, along section of roads, etc.). In this paper a novel approach to modelling RLR with traffic micro-simulation demonstrated in (Bell et al., 2012) has been applied to demonstrate the transferability of the approach and to evaluate the ability of the micro-simulation model to reproduce the RLR phenomena when traffic light operations (Cycle and green timings) changes. For consistency of the results and to demonstrate the transferability of Bell et al. (2012) approach the AIMSUN model (http://www.aimsun.com/wp/) was used.

The microsimulation model has been setup to represent the 4 arm junction studied, including traffic light setup, pedestrian crossings and stop-line. Different parameters, such as reaction time at traffic light stop-line, max/min acceleration and deceleration, headway and speed limit have been included to improve the capacity of the model to reflect driver behaviour at the signalised intersection simulated. Hourly Origin-Destination matrices for each of the 11 hours monitored each day have been setup using the traffic flow measurements made during the survey period.

5. Validation and Discussion of the Micro-simulation Model of RLR

A validation exercise has been carried out in order to assess the model performance against the ability to reproduce measured data. The comparison of flows, from direct observation, with those predicted, from the microsimulation shows a very good agreement (with a regression coefficient of R^2 of 0.99, Figure 3) for the whole flow range, 110 to 1,080 vehicles per hour, measured at the junction.
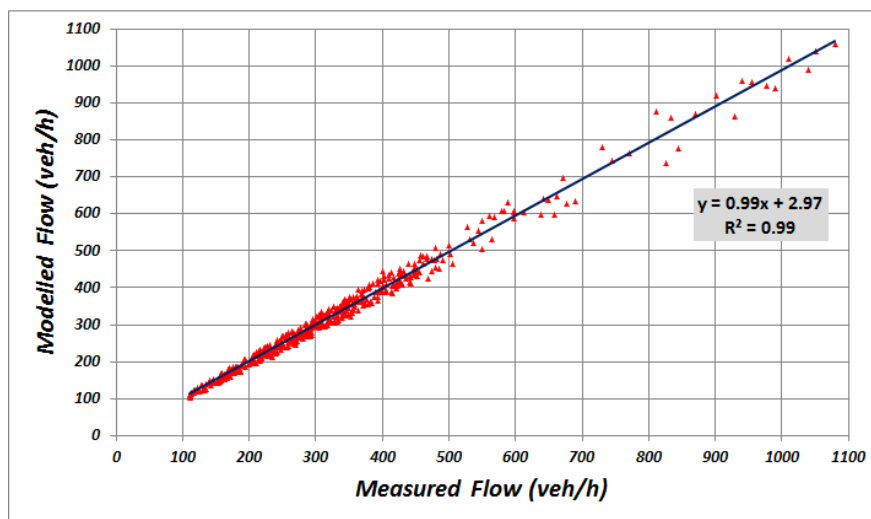

Figure 3. Validation of the simulation model used at the junction studied

It is important to mention that the model was able to reproduce to such a high level of performance the traffic at the studied junction due to the richness of data available as well as to the relatively straight forward junction geometry, and the richness of the speed and flow information, of essentially an isolated junction with fixed time plans.

In order to simulate the RLR events, and adopting the same methodology used in Bell et al. (2012), the micro simulation model was set up with a double traffic light at each entry link, one representing the existing traffic light and the second, few metres further down, used as a virtual one. The double traffic light sets per link have been set up with same cycle time and with different offsets (0 to 4 secs with 1sec steps) to ensure that as the vehicle crosses the first stop-line, it is presented with a green with 1, 2, 3 and 4 seconds delay to the start of the red to stop at the following stop-line. In this way, the two possible behaviours, decelerating to stop and accelerating to pass through the junction (in reality during the amber), equivalent to a violation, were modelled. This approach was adopted because the micro-simulator is set up not to allow vehicles to cross a stop-line during the red, thus only the unintentional RLR behaviours have been modelled. The collection of the violations data for the post-processing and analysis was made possible by locating two sets of detectors to count the number of vehicles crossing second by second, specifically for the time period after the amber starts, to evaluate the number of violation for each cycle (see Figure 4).

Figure 4. RLR microsimulation model layout in 2D (left) and 3D (right)

In order to identify the modelled violations, vehicles crossing the stop line during the amber have been extrapolated from the model outputs using a filtering algorithm.

In Figure 5 an example of how the microsimulation model represents the vehicle crossing during the amber and generating an unintentional violation.

To carry out the RLR microsimulation model analysis hourly matrices between 7:30 hrs and 22:30 hrs, for the 13 days of observations have been prepared and imported into AIMSUN.

Five runs per scenario have been produced in order to calculate the average value of the number of unintentional violations associated to the main and secondary directions.
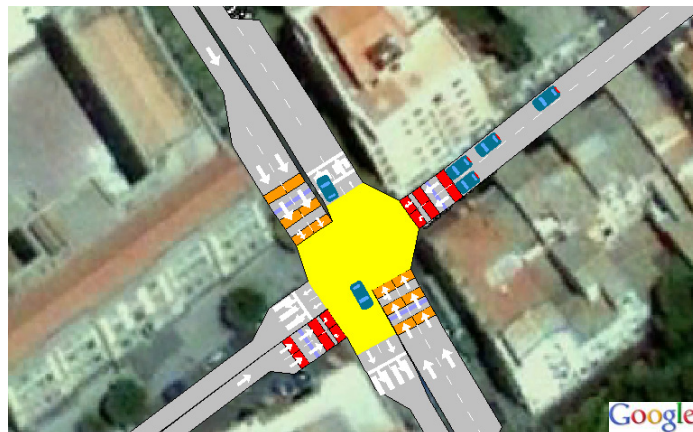

Figure 5. Unintentional RLR microsimulation

6. Comparison of RLR Measurements with Microsimulation Results

The results obtained from the microsimulation model analysis are represented and compared in Figure 6.
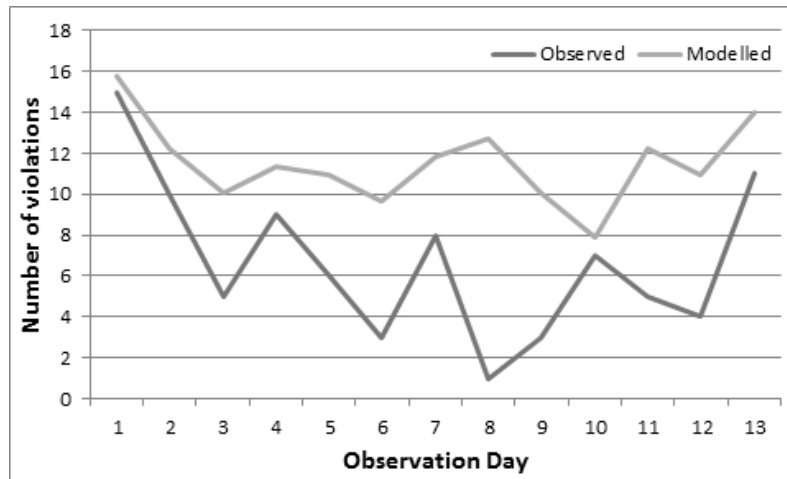

Figure 6. Observed and modelled violations during the 13 days of observation and simulation

The microsimulation performs well at the beginning and at the end of the observed period with small differences, this is probably due to the fact that microsimulation model and the representation of driver behaviour is more similar to a long-term observation with the same traffic light setup, where the usual drivers know well the timings and are more aggressive and take more risks. In fact while between the 3rd and 12th day most of the observation show a similar behaviour/trend, the simulated violations are generally greater than the observed ones. This it is probably due to the fact that during the 3rd and 8th day the traffic light setup has been changed daily generating uncertainty to the usual drivers who felt less confident and then less aggressive, leading to a substantial reduction on the number of violations with the minimum of 1 violation the 8th day. For this reason the microsimulation model results can be considered as an upper limit of the number of violation expected to occur in the junction for a certain traffic light setup and a certain traffic flow. In fact at the end of the period of observation the difference between simulated and observed violations reduce to 3.

Finally, Between the 8th and 11th day the microsimulation presents opposite trend if compared to the measurements, day 8 is the first day of the constant setup after a week of continuous changes in traffic-light cycle and green, this could be the reason of a specific and very different behaviour from the drivers that after day 11th restart to match the simulation trend.

The information collected during the 13 days of survey highlighted that the variability of the traffic light settings sensibly influenced the number of violation rates by reducing the total numbers, this suggests that a possible countermeasure could be the implementation of an actuated traffic light that accommodating the variations of traffic flows during the day will modify the cycle and green timings generating a limited but effective variability that could replicate the changes implemented during the 13 days of survey. This is expected to produce the beneficial effects measured. Moreover, looking in detail at the junction layout and on-street parking characteristics (Figure 7), another potential countermeasure could be to implement the prohibition of on-street parking in close proximity of the junction, in particular on the right hand side where the traffic signal is located (Figure 7), this will improve significantly the visibility of the traffic signal at a sufficient distance and will capture the drivers' attention with positive effects by reducing the unintentional violations.

Figure 7. View of the junction studied from Viale Diaz

7. Conclusions

Building on the findings from previous study from the authors (Bell et al., 2012), in this paper, a detailed analysis of the violations occurred in a 4 arm junction in Enna has been carried out and interesting patterns of the evolution and variation of the number of violation varying the traffic light setup of the junction have been identified.

By testing the new modelling framework using the signalised junction in Enna the first step of a more comprehensive transferability exercise of the modelling approach has been addressed in this paper, also the variability of the traffic light settings allows to make more robust the modelling framework. The comparison between the modelled and observed violation confirms that, except between the 8th and 11th day, the microsimulation model performs well in predicting the number of unintentional RLR violations, while when changing daily the timings/cycle the number of observed violations is lower than the modelled, but it follows the same trend.

However, between the 8th and 11th day, this preliminary study highlighted a different prediction of RLR violations from the model this, associated to the limited number of days of survey, in particular from day 8 when a constant traffic-light setup has been implemented after a week of modification, clearly identifies an area where further research is required.

These results although encouraging and broadly consistent with the previous study, suggest the need of further investigation on the adaptability of the drivers to the different traffic-light setup, in particular it is anticipated that a more comprehensive survey covering a longer period of observation with limited variation of traffic light timings through each survey period will be developed as part of a future study.

Moreover, while this paper provides other evidence of the potential of the methodological approach developed in Bell et al. (2012) it is intention of the authors to continue to test and diversify the case studies in future research in order to make more robust and consistent the findings and the method adopted.

Finally, because of the potential value for traffic operators and Local Authorities of the extended use of existing micro-simulation models it is needed and planned by the authors to carry out a comprehensive assessment of the entire area of an existing micro-simulation, in order to explore the potential of this approach to identify relationships between violation rates and layout, signal settings and traffic flow volumes of the signalised junctions within the entire area.

References

Allos, A. E., & Al Haidithi, M. I. (1992). Driver Behavior During Onset of Amber at Signalized Junctions. *Traffic Engineering and Control, 33*(5), 312-317.

Bell, M. C., Galatioto, F., Giuffrè, T., & Tesoriere, G. (2010). An extended analysis of red-light violation model using traffic micro-simulation. proceeding 8th Malaysian Road Conference & Exhibition 2010, October 2010, Kuala Lumpur (Malaysia).

Bell, M. C., Galatioto, F., Giuffrè, T., & Tesoriere, G. (2012). Novel application of red-light runner proneness theory within traffic microsimulation to an actual signal junction. *Accident Analysis & Prevention, 46*, 26-36. http://dx.doi.org/10.1016/j.aap.2011.12.010

Bonneson, J., Brewer, M., & Zimmerman, K. (2001). *Review and Evaluation of Factors that Affect the*

Frequency of Red Light-Running. FHWA/TX-02/4027-1, Washington, DC: Federal Highway Administration.

Campbell, B. N., Smith, J. D., & Najm, W. G. (2004). *Analysis of Fatal Crashes Due to Signal and Stop Sign Violations*. Washington D.C. (USA), Department of Transportation DOT HS 809 779.

CEC (Commission of the European Communities). (2008). *Proposal for a Directive of the European Parliament and of the Council Facilitating Cross-boarder Enforcement in the Field of Road Safety*, COM (2008) 0151, Brussels.

Chodur, J., Ostrowski, K., & Tracz, M. (2011). Impact of Saturation Flow Changes on Performance of Traffic Lanes at Signalised Intersections. *Procedia - Social and Behavioral Sciences, 16*, 600-611. http://dx.doi.org/10.1016/j.sbspro.2011.04.480

Elmitiny, N., Yan, X., Radwan, E., Russo, C., & Nashara, D. (2010). Classification analysis of driver's stop/go decision and red-light running violation. *Accident Analysis and Prevention, 42*, 101-111. http://dx.doi.org/10.1016/j.aap.2009.07.007

Federal Highway Administration. (2003) *Making Intersections Safer: A Toolbox of Engineering Countermeasures to Reduce Red-Light-Running. An Informational Report*. Institute of Transportation Engineers, Publication No. IR-115, ISBN: 0-935403-76-0, Washington DC, USA.

Giuffrè, O., Granà, A., Marino, R., & Corriere, F. (2011). Handling underdispersion in calibrating safety performance function at Urban, four-leg, signalized intersections. *Journal of Transportation Safety and Security, 3*(3), 174-188. http://dx.doi.org/10.1080/19439962.2011.599014

Giuffrè, T., & Rinelli, S. (2006). Evaluation of proneness to red light violation. A quantitative approach. *Transportation Research Record, 1969*, 112-119. Retrieved from http://www.aimsun.com/wp/ (01/09/2012)

Kennedy, J., & Sexton, B. (2009). *Literature Review of Road Safety at Traffic Signals and Signalised Crossings*. Published Project Report (PPR436), Transport Research Laboratory.

Koh, P. P., & Wong, Y. D. (2007). Driving situations and driver decisions at road traffic signals. *Journal of Advanced Transportation, 41*, 53-68. http://dx.doi.org/10.1002/atr.5670410105

Milazzo, J. S., Hummer, J. E., & Prothe, L. M. (2001). *A Recommended Policy of Automated Electronic Traffic Enforcement of Red Light Running Violations in North Carolina*. Final Report, Institute for Transportation Research and Education, North Carolina State University, Raleigh, N.C.

Mohamedshah, Y. M., Chen, L. W., & Council, F. M. (2000). *Association of Selected Intersection Factors with Red-Light Running Crashes*. Publication FHWA RD-00-112. FHWA, U.S. Department of Transportation.

Retting, R. A., & Williams, A. F. (1996). Characteristics of Red Light Violators: Results of a Field Investigation. *Journal of Safety Research, 27*(1), 9-15.

Sze, N. N., Wong, S. C., Pei, X., Choi, P. W., & Lo, Y. K. (2011). *I*s a combined enforcement and penalty strategy effective in combating red light violations? An aggregate model of violation behavior in Hong Kong. *Accident Analysis & Prevention, 43*(1), 265-271. http:// dx.doi.org/10.1016/j.aap.2010.08.020

Wissinger, L. M. (2000). *Issues Surrounding the Operation and Installation of Red Light Cameras. Master's thesis*. N. C. State University.

A Proposed Centrality Measure:
The Case of Stocks Traded at Bursa Malaysia

Shamshuritawati Sharif[1,2] & Maman Abdurachman Djauhari[1]

[1] Faculty of Science, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia

[2] College Arts and Sciences, Universiti Utara Malaysia, Sintok, Kedah, Malaysia

Correspondence: Shamshuritawati Sharif, Faculty of Science, Universiti Teknologi Malaysia, 83100 UTM Skudai, Johor, Malaysia; College Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia. Tel: 60-1-9424-8001. E-mail: shamshurita@uum.edu.my

Abstract

In this paper we propose the average of weights of all links adjacent to each stock as a centrality measure. This measure, besides the traditional centrality measures such as degree centrality, betwenness centrality, closeness centrality and eigenvector centrality will be helpful in interpreting the network topology of stocks markets. A case study of 90 stocks market traded at Bursa Malaysia will be presented and discussed to illustrate the advantage of the proposed measure.

Keywords: correlation matrix, distance matrix, Kruskal algorithm, minimum spanning tree, sub-dominant ultrametric

1. Introduction

The stock market has become an increasingly significant subject of the economy and is used by tens of thousands of companies to access equity capital, and tens of millions of investors to pursue opportunities around the world. The behaviour of stock prices is a subject of enduring interest to the investors, policymakers, and economists, and it is widely believed to be the predictor of economic activity (Peek & Rosengren, 1988; Muradoglu et al., 2000). Now days, the researchers from various disciplines i.e. mathematics, financial analysis (Daly & Fayyad, 2011; Da Costa et al., 2005; Zhang et al., 2009) and theoretical physics (Eom et al., 2009; Eom et al., 2010) are given their attention on analyzing the stock behaviour. In general, the behaviour of a stock market will be influenced by the behaviour of others stocks traded in that market. Mathematically, the interrelationships among stocks are customarily represented by the correlations among the logarithm of stock returns. The correlation structure, together with the corresponding stocks, constitutes a complex system in the form of a network. Recently, the network topology of stocks has been introduced in the field of econophysics to understand the interaction among the stocks. Mantegna (1999) presents the complicated relationship between stocks in a topological space by using a visualization mechanism, specifically minimum spanning tree. Since the work of Mantegna, subsequently many studies have confirmed the various properties of stock networks constructed.

To the aid of interpretation, centrality measure can help us to have a better understanding about the information contained in the network as well as to enrich the economic interpretation. The role of importance of each particular stock can be express by the use of centrality measure such as degree centrality, closeness centrality between centrality and eigenvector centrality. However, the problem with those measures is that they take into account only the number of direct and indirect links between two stocks. For that reason, we acknowledge the weight between two different stocks to propose a new measure of centrality which represents the "power" of stock. Later on, to illustrate the role of the propose measure, in this paper we conduct a study on the daily stock price data of 90 stocks traded at Bursa Malaysia from January 1, 2007 until December 31, 2009. For that purpose, those stocks will be viewed as a complex system consisting of 90 stocks as nodes connected by [90-1]*[90]/2=4005 links each of which corresponds to the correlation coefficient between two different nodes.

The nodes and links will be considered as a network or, more specifically, a weighted undirected graph (Jayawant & Glavin, 2009). This point of view is useful in order to visualize, simplify, and summarize the most important information contained in that complex system. The rest of paper is organized as follows. In Section 2

we discuss the analysis of that complex system by using the propose centrality measure as well as the current existing measures. To illustrate the advantage of the proposed measure, a case study on stock market will be presented and discussed in Section 3. At the end of this paper, we will draw attention to a conclusion.

2. Proposed Centrality Measure

In this section we discuss the analysis of a complex system by using network topology approach. This allows us to visualize and simplify a complex system, and to summarize the information contained therein. We show that the centrality measures usually used as the principal tools to summarize the information are not sufficient. This motivates us to propose another measure.

2.1 Network Topology

The essence of a network is its nodes (stocks) and the way how they are linked. Network analysis was originally developed in computer science (De Nooy et al., 2004). Nowadays, it has been used in various fields of study. See, for example, Krichel and Bakkalbasi (2006) in sociology, Mantegna (1999) and Miccichè et al. (2003) in finance, and Park and Yilmaz (2010) in transportation.

In financial industry, network analysis starts with correlation matrix followed by transforming it into a distance matrix (Mantegna & Stanley, 2000). From this matrix we construct a minimum spanning tree (MST) and the corresponding sub-dominant ultrametric (SDU) distance matrix. For this purpose we use Kruskal algorithm (Kruskal, 1956) as suggested in Mantegna and Stanley (2000) and Jayawant and Glavin (2009). MST will then be used to construct network topology of stocks.

The Kruskal algorithm is a graph without a cycle that connects all nodes with links. The correlation coefficient can vary between $-1 < \rho_{ij} < +1$ whereas the distance can vary between $0 < d_{ij} < 2$. Here, small values of the distance imply strong correlations between stocks. This is a simplification of the complex system of stocks and their correlation structure which will be used to summarize the most important information.

The visualization of MST can be made possible by using the open source called 'Pajek' (Batagelj & Mrvar, 2003; Batagelj & Mrvar, 2011; De Nooy et al., 2004). Furthermore, to interpret the MST we use the standard tools, i.e., centrality measures. To make the network topology more attractive and easy to interpret, we use the Kamada Kawai procedure provided in Pajek (Kamada & Kawai, 1989).

2.1.1 Correlation Matrix

The stock network visually displays the significant p-1 links among all possible links, $(p$-1)* p/2. This is based on the correlation matrix between stocks, using the MST method. The MST, a theoretical concept in graph theory (West, 2000), is also known as the single linkage method of cluster analysis in multivariate statistics (Gower & Ross, 1969; Everitt, 1980).

Let $P_i(t)$ be the stock price of a stock i and $R_i(t)$ be the logarithm of daily stock return at day t in a given period, defined as:

$$R_i(t) = \ln P_i(t+1) - \ln P_i(t).$$ (1)

for all i = 1, 2, … , 90. Equation (1) defines a complex system among stocks in the form of stock networks. To filter the information contained therein, we construct a correlation matrix among those stocks, is a symmetric matrix of size 90×90 where the element in the i-th row and j-th column is,

$$\rho_{ij} = \frac{\langle R_i R_j \rangle - \langle R_i \rangle \langle R_j \rangle}{\sqrt{\left(\langle R_i^2 \rangle - \langle R_i \rangle^2\right)\left(\langle R_j^2 \rangle - \langle R_j \rangle^2\right)}}$$ (2)

representing the correlation coefficient between i-th and j-th stocks (Mantegna & Stanley, 2000). That correlation coefficient quantifies the degree of linear relationship between i-th and j-th stocks. By definition, $\rho_{ii} = 1$ for all i and ρ_{ij} can vary from -1 to 1 for all $i \neq j$ where,

$$\rho_{ij} = \begin{cases} 1 & \text{means perfectly positive linear relationship} \\ 0 & \text{means no linear relationship} \\ -1 & \text{means perfectly negative linear relationship} \end{cases}$$

2.1.2 Distance Matrix

To analyze the network, we transform the correlation matrix into a distance matrix by using the following formula (Mantegna & Stanley, 2000).

$$d_{ij} = \sqrt{2(1-\rho_{ij})} \tag{3}$$

This d_{ij} is a distance between the i-th and j-th stocks since it satisfies the following three properties; (i) $d_{ij} \geq 0$ and $d_{ij} = 0 \Leftrightarrow X_i = X_j$, (ii) $d_{ij} = d_{ji}$, and (iii) $d_{ij} \leq d_{ik} + d_{kj}$. The first property tells us that two stocks that are perfectly correlated (either positive or negative), $|\rho_{ij}| = 1$, will be represented by a single point in Euclidean space ($d_{ij} = 0$). Moreover, $0 \leq d_{ij} \leq 2$.

The second property is symmetric property; the distance between the i-th and j-th stocks is equal to the distance between the j-th and i-th stocks. In other words, the correlation between the i-th and j-th stocks is equal to the correlation between the j-th and i-th stocks ($\rho_{ij} = \rho_{ji} \Leftrightarrow d_{ij} = d_{ji}$).

The last property is well known as triangular property. From (2), we conclude that, in general, the higher the correlation coefficient the smaller the distance.

By using Equation (3), we obtain a distance matrix D of size 90×90 with d_{ij} as the element in the i-th row and j-th column. It is this matrix that we analyze in the rest of the paper.

2.1.3 Kruskal's Algorithm

A spanning tree is a subset of a graph which has no cycles and includes all of the nodes of the original graph, but usually not all the links. A minimum spanning tree is a spanning tree that has smaller (i.e minimum) sum of the weights of its links than any other spanning tree. When the links are weighted where their weights in this paper is representing by the distance, the problem is to find the tree that has minimal distance. In this case, Kruskal's algorithm can help to solve the problem to find the tree that has minimal distance. This is a simple algorithm since the links are selected and included to the tree in increasing order of their weights. But, we have to stop it when it create a cycle or looping.

2.1.4 Information Summarization

To visualize, simplify and summarize the important information contained in the network represented by D, we use the notion MST as discussed in Mantegna and Stanley (2000). Then, we determine MST by using Kruskal algorithm (Kruskal, 1956).

2.2 Centrality Measures

From network analysis view point, the role or degree of importance of each particular node can be analyzed by using its centrality measures such as degree, betweenness, and closeness centralities. These will help us to find the most important nodes in the network structure (Xu et al., 2009; Abbasi & Altmann, 2010; Monárrez-Espino & Caballero-Hoyos, 2010).

Degree centrality indicates the connectivity of nodes. It provides information on how many other nodes are connected with a particular node. On the other hand, betweenness centrality is reflects the extent to which a node lies in relative position with respect to the others (Freeman, 1977). This measure indicates the potentiality of node to influence the others. Closeness centrality measures how close a node is to all other nodes in terms of correlations. Closeness can also be regarded as a measure of how long the information is to spread from a given node to other reachable nodes. Nevertheless, eigenvector centrality, a point centrality measure introduce by Bonacich in 1972. The key idea is to express that an important node is connected to important neighbours (other nodes).

Those measures are computed based on the MST as follows (Borgatti, 1995; Sieczka & Holyst, 2009; Park & Yilmaz, 2010):

(i) Degree centrality of node i is $d_i = \sum_{j=1}^{n} a_{ij}$ where $a_{ij} = 1$ if the i-th and j-th nodes are linked and 0 otherwise.

(ii) Betweenness centrality of node i, b_i, is the ratio of the number of path passing through i between two different nodes and the number of all possible paths from j to k for all j and k where $j \neq i$ and $k \neq i$.

(iii) Closeness centrality of node i, c_i, is the ratio of the number of links in the MST, which is equal to (n-1), and the number of links in the path from i to j for all $j \neq i$.

(iv) Eigenvector centrality of node i is, $ev_i = \lambda^{-1} \sum_{j=1}^{n} a_{ij} e_j$ where $(e_1, e_2, ..., e_n)^t$ is the eigenvector of A that corresponds to the largest eigenvalue λ.

Degree centrality is the simplest of the node centrality measures by using the local structure around nodes only. In order to identify the role of importance, degree centrality is no longer appropriate to be the best measure. The higher the degree centrality does not reflect to the strength of each particular node.

Due to that limitation of degree centrality, in this subsection we introduce "average of weights" as another measure. It is the average of weights of all links adjacent to each node. This measure reflects the strength of influence of a particular node to the others. Thus, the larger the scores represent the powerful of that particular stock. In all measures, the node that has larger scores is considered to be more central in terms of it influence to the others.

3. A Case Study

We utilized 90 stocks that were traded on the stock market at Bursa Malaysia. The individual stocks that posted daily prices for the last 3 years from January 1, 2007 to December 31, 2009. The 1096 daily data can be retrived from Bloomberg Professional®. Based on the MST issued from Matlab version 7.8.0 (R2009a), we present the top 15 scores of centrality measure discussed previously in Table 1 and Table 2.

Table 1. The top 15 scores of degree and betweeness centrality

Top	Degree		Betweeness	
1	Genting Bhd	7	Wah Seong Corp	0.708
2	Wah Seong Corp	6	Genting Bhd	0.594
3	MMC Corp Bhd	5	YTL Cement Bhd	0.495
4	Kuala Lumpur Kep	4	MMC Corp Bhd	0.415
5	Genting Plantation	4	Parkson Holding	0.357
6	IJM Land Bhd	4	WCT Bhd	0.335
7	Malaysian Res Co	4	CIMB Group Banking	0.331
8	WCT Bhd	4	AMMB Holding Bhd	0.308
9	DRB-HICOM Bhd	4	IOI Corp Bhd	0.287
10	Supermax Corp	4	IJM Corp Bhd	0.261
11	CIMB Group Banking	3	Genting Plantation	0.260
12	IOI Corp Bhd	3	Supermax Corp	0.241
13	AMMB Holding Bhd	3	RHB Capital Bhd	0.204
14	RHB Capital Bhd	3	Kuala Lumpur Kep	0.188
15	IJM Corp Bhd	3	Hong Leong Bank	0.184

Table 2. The top 15 scores of closeness and eigenvector centrality

Top	Closeness		Eigenvector	
1	Wah Seong Corp	0.200	Wah Seong Corp	0.500
2	YTL Cement Bhd	0.190	Genting Bhd	0.427
3	Genting Bhd	0.184	MMC Corp Bhd	0.296
4	MMC Corp Bhd	0.179	Genting Bhd	0.292
5	RHB Capital Bhd	0.170	RHB Capital Bhd	0.211
6	Sunway City Bhd	0.169	IJM Corp Bhd	0.208
7	CIMB Group Banking	0.166	CIMB Group Banking	0.198
8	Tanjong PLC	0.164	Sunway City Bhd	0.181
9	Dialog Group Bhd	0.164	Dialog Group Bhd	0.167
10	IJM Corp Bhd	0.163	Tanjong PLC	0.167

11	*Parkson Holding*	0.163	*EON Capital Bhd*	0.143
12	*EON Capital Bhd*	0.156	*Multi-Purpose*	0.143
13	*AFFIN Holding*	0.156	*UNISEM(M) Bhd*	0.143
14	*Multi-Purpose*	0.156	*AFFIN Holding*	0.143
15	*UNISEM(M) Bhd*	0.156	*Parkson Holding*	0.141

However, by using Pajek, we can visualize the interrelationship among 90 stocks with more attractively. Therefore, in Figure 1 - Figure 4 we present their network topology with respect to their centrality measure. The size and colour of the node represent the score of centrality measure and the rank of importance for degree centrality, betwenness centrality, closeness centrality, eigenvector centrality and average of weight centrality.

From Figure 1, degree centrality measure, we learn that, the highest number of links in the network belongs to *7-Genting Bhd* (red point). It is followed by *83-Wah Seong Corp* (blue point) and *25-MMC Corp Bhd* (yellow point) are 6 and 5 links, respectively. Each of the fololowing has 4 links: *12-Kuala Lumpur Kep, 34-Genting Plantation, 61-IJM Land Bhd, 64-Malaysian Res Co, 66-WCT Bhd, 67-DRB-HICOM Bhd, 72-Supermax Corp* (green points). The rests are of 1, 2 and 3 links only. The higher the number of links is the higher the influence of that particular stock to the others.

According to the betweenness centrality, see Figure 2, the most important nodes is *83-Wah Seong Corp* (red point). It has an excellent position compared to the others where the information flow in the network can easily reach others in the network followed by, in order of importance: *7-Genting Bhd* (blue point or the second most important), *73-YTL Cement Bhd* and *25-MMC Corp Bhd* (yellow points or the third most important). This means that those stocks strongly dominance the other stocks, especially the neighbour which is closely to their corner. As example, if come out any shift in the price of *7-Genting Bhd,* the price of the following stocks: *73-YTL Cement Bhd*, *39-Affin Holding*, *38-EON Capital Bhd*, *2-CIMB Group Bhd*, *28-IJM Corp Bhd*, *65-Multi-Purpose*, and *89-UNISEM(M) Bhd* will directly get the impact.
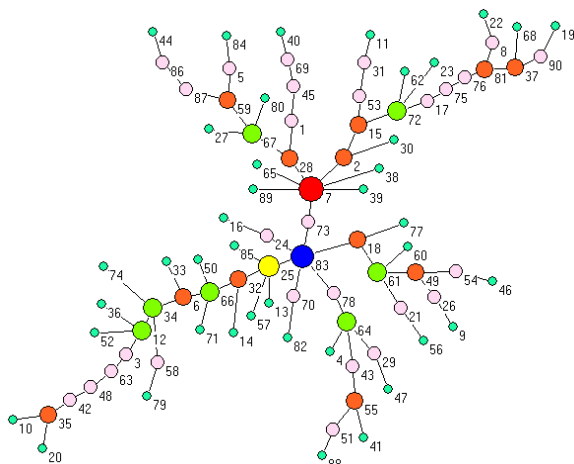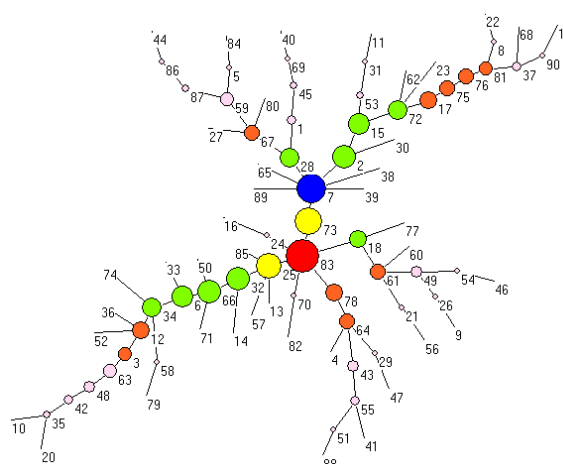

Figure 1. Degree centrality   Figure 2. Betwenness centrality

In Figure 3, we present the closeness centrality. The key player in this analysis is represented by *83-Wah Seong Corp* (red point). It plays the most important role in the network and this node is the closest node to the others. The second closest nodes to the others are *7-Genting Bhd*, *73-YTL Cement Bhd* and *25-MMC Corp Bhd* (blue points). The third closest is yellow points; *18-RHB Capital Bhd, 78-Sunway City Bhd, -CIMB Group Banking, 24-Tanjong PLC, 70-Dialog Group Bhd, 28-IJM Corp Bhd, 32-Parkson Holding, 38-EON Capital Bhd, 39-AFFIN Holding, 65-Multi-Purpose, 89-UNISEM(M) Bhd, 13-Genting Malaysia, 57-Star Publication, 85-Lingkaran Trans, 15-AMMB Holding Bhd.*

Based on eigenvector centrality, see Figure 4, *83-Wah Seong Corp* (red point) has the highest scores of centrality measure in the network. This result is similar to betwenness centrality and closeness centrality. The blue points or the second most important are represented by *7-Genting Bhd* followed by third most important stocks; *73-YTL Cement Bhd, 25-MMC Corp Bhd, 18-RHB Capital Bhd,* and *28-IJM Corp Bhd* (yellow points).
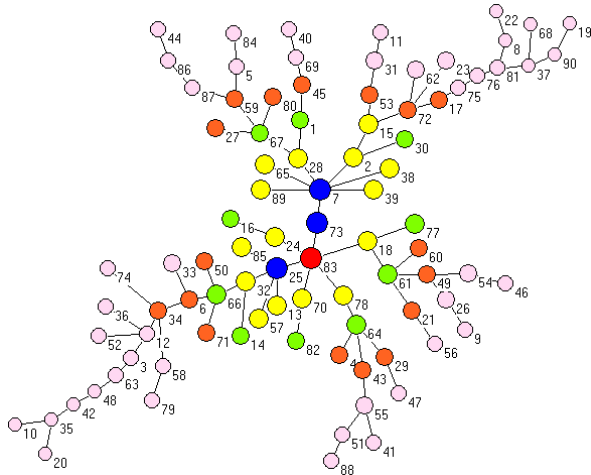
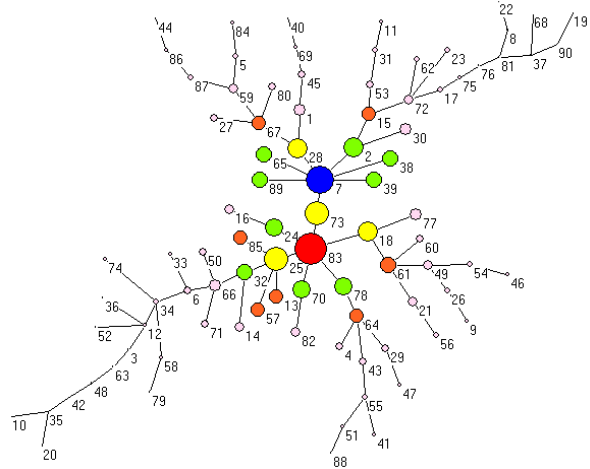Figure 3. Closeness centrality   Figure 4. Eigenvector centrality

Average of weight centrality can be used to indicate the average correlations between a particular node and the other nodes adjacent to it. In terms of degree, *7-Genting Bhd* are the most dominance stocks while in terms of average of weights (see Table 3 and Figure 5) the most influential is *20-Telekom* (red point), followed by *19-Brit Amer Tobacc* (blue point) as the second important, and *90-JT International* (yellow point) as the third important. The fourth important stock is *68-KFC Holdings* (green point) followed by the orange points; *33-Berjaya Sports, 35-Berjaya Land, 37- & Neave, 10-DIGI.com, 57-Star Publication, 31-Malaysian Airport, 11-Plus Expressway, 79-NCB Holdings, 50-SHELL Refining,* and *14-YTL Power*. The rest are having small average of weights (pink points).

Table 3. The top 15 scores of average of weights centrality

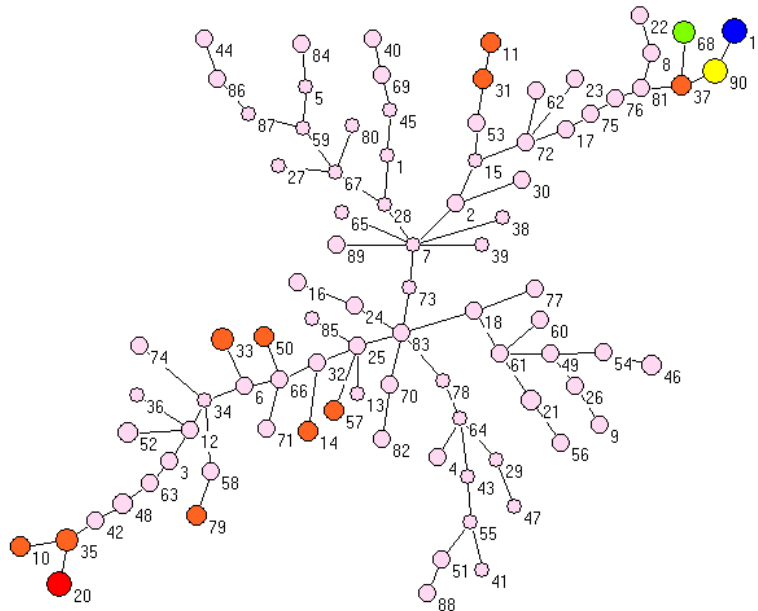| Top	Average of Weights	
1	Telekom	1.065
2	Brit Amer Tobacc	1.026
3	JT International	0.925
4	KFC Holdings	0.815
5	Berjaya Sports	0.790
6	Berjaya Land	0.739
7	Fraser & Neave	0.701
8	DIGI.com	0.662
9	Star Publication	0.658
10	Malaysian Aiport	0.651
11	Plus Expressway	0.646
12	NCB Holdings	0.618
13	SHELL Refining	0.607
14	YTL Power	0.600
15	Boustead Holding	0.595

Figure 5. Average of weights centrality

Figure 1 and Figure 5, represent an MST with degree centrality and average of weights, respectively. If degree centrality refer to the number of links adjacent to a node, average of weights is the sum of those links' weight divided by the number of links. The latter measure represent the average influence given by a node to the others adjacent to it. From Table 1, we see its advantage compared to the former and learn that those measures are different.

4. Conclusion

We have introduced "average of weights" as a new centrality measure. The advantage of proposed centrality measure is by considering the weight instead of number of links only. Its advantage is illustrated by using 90 stocks market traded at Bursa Malaysia, together with the other established centrality measures, and could help to enrich the economic interpretation of network topology.

According to the five centrality measures, after using the Pareto analysis based on the top 15 stocks of highest scores in each centrality measures (Table 1), the following six stocks are the most in influencing stocks market; *CIMB Group Banking, Genting Bhd, RHB Capital Bhd, Wah Seong Corp, MMC Corp Bhd* and *IJM Corp Bhd*. These six stocks should be paid more attention by the investors in order to make the investment.

For further research, we render two potential researches. First, we can perform a comparison study on the proposed measure by using difference algorithm for constructing a minimum spanning tree. Furthermore*,* its applicability to other networks can be performed. Second, we can investigate its mean evolution distribution as well as its variance when involving more than one correlation structure. Many more research idea can be carried out to enrich this topic.

Acknowledgment

References

Abbasi, A., & Altmann, J. (2010). *On the Correlation between Research Performance and Social Network Analysis Measures Applied to Research Collaboration Networks*. TEMEP Discussion Paper, No. 2010, Seoul National University, Korea.

Batagelj, V., & Mrvar, A. (2003). *A Density based approaches to network analysis: Analysis of Reuters terror news network*. Washington, D. C.: Ninth Annual ACM SIGKDD.

Batagelj, V., & Mrvar, A. (2011). PAJEK: Program for Analysis and Visualization of Large Networks, version 2.02. Retrieved from http://pajek.imfm.si/doku.php?id=download

Borgatti, S. P. (1995). Centrality and AIDS. *Connections, 18*(1), 112-114.

Da Costa, Jr. N., Nunes, S., Ceretta, P., & Da Silva, S. (2005). Stock-market co-movements revisited. *Economics Bulletin, 7*(3), 1-9.

Daly, K. J., & Fayyad, A. (2011). Can Oil Prices Predict Stock Market Returns?. *Modern Applied Science, 5*(6), 44-54. http://dx.doi.org/10.5539/mas.v5n6p44

De Nooy, M., Mvrar, A., & Batagelj, V. (2004). *Exploratory Social Network Analysis with Pajek*. Cambridge: Cambridge University Press.

Eom, *C.*, Oh, G., Jung, W. S., Jeong, H., & Kim, S. (2009*). Topological properties of stock networks based on minimal spanning tree and random matrix theory in financial time. Physica A, 388*, 900-906. http://dx.doi.org/10.1016/j.physa.2008.12.006

Eom, C., Oh, G., Jung, W. S., & Kim, S. (2010). The effect of a market factor on information flow between stocks using the minimal spanning tree. *Physica A, 389*, 1643-1652. http://dx.doi.org/10.1016/j.physa.2009.12.044

Everitt, B. S. (1980). *Cluster Analysis (2nd ed.)*. London: Heineman Educational Books.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*(1), 35-41. http://dx.doi.org/10.2307/3033543

Górski, A. Z., Drożdż, S., Kwapień, J., & Oświęcimka, P. (2008). Minimal Spanning Tree Graphs and Power Like Scaling in FOREX Networks. *Acta Physica Polonica A, 114*(3), 531-538.

Gower, J. C., & Ross, G. J. S. (1969). Minimum. Spanning trees and single-linkage cluster analysis. *Applied Statistic, 18*(1), 54-64. http://dx.doi.org/10.2307/2346439

Jayawant, P., & Glavin, K. (2009). Minimum spanning trees. *Involve a journal of mathematics, 2*(4), 439-450. http://dx.doi.org/10.2140/involve.2009.2.439

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters (Elsevier), 31*(1), 7-15. http://dx.doi.org/10.1016/0020-0190(89)90102-6

Krichel, T., & Bakkalbasi, N. (2006). *A Social Network Analysis of Research Collaboration in the Economics Community*. The International Workshop on Webometrics, Informetrics and Scientometrics & Seventh COLLNET Meeting, France.

Kruskal, J. B. (1956). On the shortest spanning subtree and the travelling salesman problem. *Proceedings of the American Mathematical Society, 7*(1), 48-50. http://dx.doi.org/10.1090/S0002-9939-1956-0078686-7

Mantegna, R. N. (1999). Hierarchical Structure in Financial Markets. *European Physical Journal B, 11*, 193-197. http://dx.doi.org/10.1007/s100510050929

Mantegna, R. N., & Stanley, H. E. (2000). *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge UK: Cambridge University Press.

Miccichè, S., Bonanno, G., Lillo, F., & Mantegna, R. N. (2003). Degree stability of a minimum spanning tree of price return and volatility. *Physica A, 324*, 66-73. http://dx.doi.org/10.1016/S0378-4371(03)00002-5

Monárrez-Espino, J., & Caballero-Hoyos, J. R. (2010). Stability of Centrality Measures in Social Network Analyses to Identify Long-Lasting Leaders from an Indigenous Boarding School of Northern Mexico. *Estudios sobre las Culturas Contemporaneas*, *16*(32), 155-171.

Muradoglu, G., Taskin, F., & Bigan, I. (2000). Causality between Stock Returns and Macroeconomic Variables in Emerging Markets. *Russian and East European Finance and Trade, 36*(6), 33-53.

Park, K., & Yilmaz, A. (2010). *A Social Network Analysis Approach to Analyze Road Networks*. ASPRS Annual Conference. San Diego, CA.

Peek, J., & Rosengren, E. S. (1988). The Stock Market and Economic Activity. *New England Economic Review,* 39-50.

Sieczka, P., & Holyst, J. A. (2009). Correlations in commodity markets. *Physica A, 388*, 1621-1630. http://dx.doi.org/10.1016/j.physa.2009.01.004

West, D. B. (2000). *Introduction to Graph Theory (2nd ed.)*. Englewood Cliffs, NJ: Prentice-Hall.

Xu, Y., Ma, J., Sun, Y., Hao, J., Sun, Y., & Zhao, Y. (2009). *Using Social Network Analysis As A Strategy For E-Commerce Recommendation*. Pacific Asia Conference on Information Systems (PACIS), India.

Zhang, J., Shan, R., & Su, W. (2009). Applying Time Series Analysis Builds Stock Price Forecast Model. *Modern Applied Science, 3*(5), 152-157.

Comparing Performances of Turbo-roundabouts and Double-lane Roundabouts

Orazio Giuffrè[1], Anna Granà[1] & Sergio Marino[1]

[1] Department of Civil, Environmental, Aerospace, Materials Engineering, University of Palermo, Palermo, Italy

Correspondence: Anna Granà, Department of Civil, Environmental, Aerospace, Materials Engineering, Viale delle Scienze, Ed. 8, Palermo 90128, Italy. Tel: 39-91-2389-9718. E-mail: anna.grana@unipa.it

Abstract

Starting from assumptions regarding the arrival process of circulating streams and according to models based on the gap-acceptance theory, the paper is aimed at comparing operational performances between basic turbo-roundabouts and double-lane roundabouts. The paper proposes applications of the Hagring model for entry capacity estimations at double-lane roundabouts and turbo-roundabouts, these latter, in particular, featured by movements with only one or two conflicting traffic streams. This model allows to use, in fact, a bunched exponential distribution to quantify the distribution of major vehicle headways; it also considers specific values different by each lane for behavioural parameters, minimum headway and conflicting traffic flow on circulating lanes.

The results obtained for the two cases examined, although influenced by the underlying assumptions, especially with regard to user behaviour at turbo-roundabouts, can give information about the convenience in choosing, at a design level, a basic turbo-roundabout rather than a double-lane roundabout. The comparison developed in this paper, indeed, can be helpful in selecting the type of roundabout and in particular in evaluating performance benefits that are obtainable from the conversion of an existing double-lane roundabout to a turbo-roundabout with similar footprint of space.

Keywords: turbo-roundabout, traditional roundabout, operating performaces

1. Introduction

1.1 Introducing the Problem

Turbo-roundabouts represent a new type of circular intersection which were designed to improve safety performances at modern roundabouts, already widely spread in the world, without compromising their efficiency. The turbo-roundabout is a specific kind of spiralling roundabout developed in The Netherlands by Fortuijn in the late 1990's. Fortuijn developed turbo-roundabouts in an attempt to deal with the drawbacks of double-lane roundabouts: while double-lane roundabouts have a higher capacity than single-lane roundabouts, they have the disadvantage of a higher driving speed through the roundabout and lane changing on the ring, hence raising the crash risk. Turbo-roundabouts were, indeed, introduced to deal with the entering and exiting conflicts occurring at double-lane roundabouts; these conflicts are eliminated at turbo-roundabouts by directing drivers to the correct lanes before entering the intersection and introducing spiral lines that guide them to the correct exit. On design principles and geometric elements of a turbo roundabout, as well as different variants of the turbo-roundabout progressively introduced in The Netherlands, can be seen e.g. Fortuijn (2009a). Other European experiences with turbo-roundabouts are referred by Brilon (2008) and Tollazzi et al. (2001).

An exhaustive evaluation of safety performances at turbo-roundabouts is not yet available because turbo-roundabout installations are still recent. It follows that the design choice between a standard double-lane roundabout or a basic turbo-roundabout can be carried out through convenience evaluations in terms of operating performances. Operating performance evaluations at turbo-roundabouts can be more complicated than roundabouts. It is should be specified that, although in both circular intersections entering vehicles must give priority to circulating vehicles, drivers before entering the turbo-rundabout have to make necessarily the choice of their destination, being forced to enter in circulating lanes physically separated by raised lane dividers.

1.2 Relevant Concepts on Capacity Models at Roundabouts

A starting point for evaluating operational performances at roundabouts and turbo-roundabouts can be represented by capacity methods for two-way-stop-controlled intersections, where vehicles on major streams have priority and vehicles on minor streams are controlled by stop. There are two primary capacity models for describing this traffic situation and computing capacity estimates: linear or exponential empirical regression models, based on observed geometric and traffic flow parameters; analytical capacity models based on gap acceptance theory. However, capacities estimated through these models can widely differ between one model and another (see e.g. Al-Madani & Saad, 2009).

Empirical regression models are based on traffic observations surveyed during short time intervals (e.g. one-minute intervals) in oversaturated conditions; then a linear or exponential regression equation is fitted to the data or a multivariate regression equation needs to be developed to take account of variation in the data caused by user behavior and geometric design features. In order to develop a regression model each traffic pattern and/or geometric situation have to be surveyed; for this purpose, a large number of operational data have to be collected. Nevertheless, empirical regression models may have poor transferability to other countries or at other times (see e.g. Pratelli & Al-Madani, 2011). Moreover, regression models do not facilitate the comprehension of the underlying traffic flow theory of determining and accepting gaps upon entering the intersection (Rodegerdts et al., 2007). According to gap acceptance models, drivers before entering the intersection have to choice an acceptable gap on the major stream; the minimum gap accepted by minor-stream drivers is the critical gap. It should also be noted that, when bunched vehicles moving along a major stream form a vehicular block, minor-stream drivers can enter the conflicting stream having priority, only when the gap following the last vehicle in the block is equal to or greater than the critical gap (Tanner, 1962). The driver behaviour variability makes that the critical gap is not a constant value, but is represented by a distribution of values. Moreover, estimation procedures for critical headway do not require sites with oversaturated conditions. Another behavioural parameter is the follow-up time, defined as the time headway between two consecutively entering vehicles, utilizing the same gap in major or circulating traffic flows at roundabouts; it can be directly surveyed on-field (Rodegerdts et al., 2007). The arrival headways in conflicting stream have to be evaluated for modeling gap-acceptance process. Thus capacity models founded on the gap-acceptance theory need to specify the probability distribution of headways between vehicles in the major stream. Capacity models homogeneous each other should be used by manoeuvre type, especially where intersections perform multiple turning movements.

Technical literature proposes exponential arrival headway distribution models: negative exponential distribution, shifted negative exponential distribution and shifted negative bunched exponential distribution. The latter was introduced by Cowan (1975; 1987) and was adopted by several authors; see eg Troutbeck (1990). Properties of the bunched exponential distribution, or otherwise known as Cowan's M3 headway distribution, were also explained by Luttinen (1999). Hagring (1998) derived the capacity of a minor traffic stream hampered by independent major streams (to cross or in which a minor stream has to merge), each of these latter featured by a bunched exponential distribution. This dichotomized distribution assumes that a proportion of all vehicles are free within each major stream and have a displaced exponential headway distribution; bunching models for parameter estimations were developed by several authors; in this regard, the reader is invited to consult the specialized literature on the subject.

1.3 Research Aims and Specific Objectives of the Paper

Recent technical literature has already proposed some studies aimed at comparing schemas of roundabouts with different geometric configuration or mode of operation, but with similar footprint of space. In the absence of suitable models to interpret the operation mode and, more in general, operating performances of schemas from time to time considered, models developed for similar patterns of intersection have been often used (see Giuffrè et al., 2012; Mauro & Branco, 2010; Giuffrè et al., 2008). The question also relates to turbo-roundabouts that, as anticipated, are of recent conception and realization.

In this paper Authors intended to assess operational performances of turbo-roundabouts and double-lane roundabouts. The schemes of standard turbo-roundabout and double-lane roundabout here examined to compare performances are shown in Figures 1 and 2. Greater consistency in assumptions was evidenced with reference to the arrival process of traffic major streams on the ring. Furthermore, for pursuing the above stated objective, entry capacity estimations were obtained by applying models founded on the gap acceptance theory. In order to analyze and compare operating performances between the circular intersections depicted in Figures 1 and 2, the Hagring capacity formula was applied to the schemes under examination (Hagring, 1998). It must be said that the convenience of the two types of roundabouts here considered was estimated in terms of degree-of-saturation

and mainly in terms of delay experienced by entering vehicles, considering the latter related to the level of service quality.
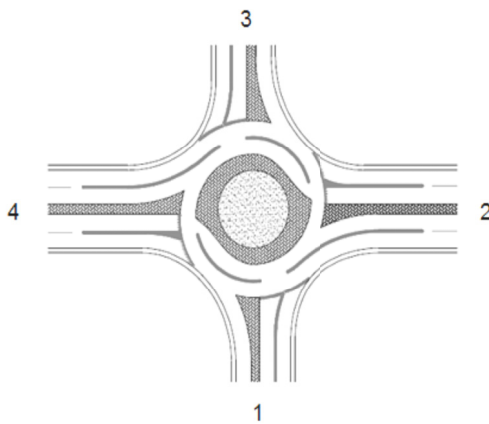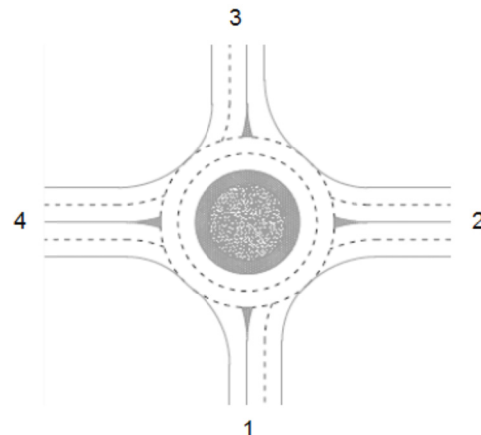

Figure 1. Basic turbo-roundabout   Figure 2. Double-lane roundabout

Description: schemes of standard basic turbo-roundabout (Figure 1) and double-lane roundabout (Figure 2) under examination; it should be noted that a priority was created for legs: legs 2 and 4 are here assumend as major entries and legs 1 and 3 as minor entries.

2. Method

In order to obtain a selection criterion for a given distribution of traffic demand, operational performances of schemes in Figures 1 and 2 were examined. The approach adopted to compare the two types of circular intersection was derived from the method developed by Mauro and Branco (2010) and also applied to compare operating performances between circular patterns of intersection, albeit different from those here studied. More specifically, the approach before quoted was adopted in order to obtain domains of convenience for the turbo-roundabout and the double-lane roundabout examined in undersaturated flow conditions and to evaluate operating advantages deriving from the choice of one or the other roundabout scheme. It should be noted that, not only this paper reports a comparison between geometric patterns of intersection different from those considered by the above cited Authors, but also different assumptions were made on conflict patterns between entering and circulating vehicles; this concerns, in particular, traffic flows faced by left turning movements from minor roads and models adopted to perfom capacity estimations. Moreover, for comparison purposes, the suitability for the selected schemes was evaluated estimating both the degree-of-saturation and the control delay. The discussion relating to the outcome of the comparison and the procedure followed for this purpose, will be preceded by a brief description of the capacity models for turbo-roundabouts and double-roundabouts, with particular reference to those applied here, with the appropriate adjustments made in relation to the specificities of the intersections under examination.

2.1 Capacity Models for Basic Turbo-roundabouts

Among capacity models for estimating entry capacity at turbo-roundabouts the linear empirical regression model proposed by Fortuijn and Harte (1997) must necessarily be mentioned. This model is based on a modification of the model derived some time earlier for entry capacity estimations at roundabouts (see Bovy et al., 1991). Fortuijn afterwards modified the factor that in Bovy model describes the effect of the circulating traffic on entry capacity, by splitting it into two parts: one for the roundabout lane with the higher volume and one for the roundabout lane with the lower volume; for a brief summary of the model, see Fortuijn (2009b). The values of aforementioned factors were then determined from observations on a turbo-roundabout built by the provincial authorities of South Holland. This model permits capacity calculations for the different legs in various types of turbo-roundabouts and delays found when traffic flow is not saturated. However, as confirmed by results of studies carried out by Brilon and Bäumer (2004), it is reasonable to conclude that an approach based on the gap acceptance theory can express the relationship between circulating traffic flow and entry capacity in a more realistic way. Mauro and Branco (2010) used the capacity model developed by Brilon et al. (1997), based on the theory of the gap acceptance; they adapted to turbo-roundabouts this model and assumed one entering lane and one circulating lane for the circular intersections that they decided to compare. A further criterion for entry capacity estimations at turbo-roundabouts was also proposed and used formulations founded on the

gap-acceptance theory for unsignalized intersections (see Giuffrè et al., 2009). Among other models to evaluate entry capacity, Hagring model (1998) is worthy of note. Hagring (1998) developed, indeed, a more general formula for capacity estimations at multi-lane intersections which takes into account behavioural and traffic flow parameters differentiated by conflicting stream; he presented a generalization of the earlier gap-acceptance models by extending Troutbeck's model (1986) to provide the expression below, rewritten to adapt it to intersection patterns under examination and in accordance with the assumptions made in this study as before stated. Thus entry lane capacity can be derived estimating capacity of a minor stream hampered by independent major streams (to cross or in which a minor stream has to merge), each featured by a shifted negative bunched exponential distribution, also referred to as Cowan's M3 headway distribution:

$$
C_e = 3600 \cdot \sum_j \frac{\varphi_j \cdot Q_{c,j}}{3600 - \Delta_j \cdot Q_{c,j}} \cdot \prod_k \left(\frac{3600 - \Delta_k \cdot Q_{c,k}}{3600} \right) \cdot \frac{exp\left[-\sum_l \frac{\varphi_l \cdot Q_{c,l}}{3600 - \Delta_l \cdot Q_{c,l}} \cdot \left(T_{c,l} - \Delta_l \right) \right]}{1 - exp\left(-\sum_m \frac{\varphi_m \cdot Q_{c,m}}{3600 - \Delta_m \cdot Q_{c,m}} \cdot T_{f,m} \right)}
$$

(1)

Although symbols in Equation 1 have the usual meaning, appropriate explanations are however opportune. It should be noted, therefore, that j, k, l, m, are indices for conflicting lanes which are repeatedly the same lanes; C_e is the entry lane capacity, in pcu/h; φ is the parameter representing the proportion of free traffic within the major stream; Q_c is the conflicting traffic flow, in pcu/h; T_c and T_f are the critical gap for circulating lane (s) and the follow-up time (s), respectively; Δ represents the minimum headway of circulating traffic (s). Thus the Hagring model (1998) resulted appropriate to evaluate entry lane capacity at turbo-roundabouts: the Hagring model allows to assume, indeed, a shifted negative bunched exponential distribution in each circulating stream along circulatory carriageway, considering values (lane-by-lane) for behavioural parameters, minimum headway and conflicting traffic flow on (one or two) lanes in the circulatory carriageway. It must also be emphasized here that at a basic turbo-roundabout vehicles entering the intersection from right and left lanes at major entries (and from right lane at minor entries) face only one antagonist traffic stream; vehicles entering the intersection from left lane at minor entries face two antagonist traffic streams.

2.1.1 Entry Capacity at Turbo-roundabouts

This section focuses on assumptions made for evaluating entry capacities at turbo-roundabout shown in Figure 1. The Hagring model in Equation 1 was specified in relation to values of conflicting traffic flow (moving on the inner circulating lane $Q_{c,i}$ or the outer circulating lane $Q_{c,e}$) faced by subject entry approach drivers, and to T_c, T_f and Δ values. For turbo-roundabout, the values based on an empiric research on turbo-roundabouts installed in the Netherlands were used (see Fortuijn, 2009b); collected values for critical gap and follow-up time were differentiated by entry and by entering lane. According to Fortuijn (2009b) only for the left entering lane at minor entries (entries 1-3 in Figure 1) two critical gap values (one for the inner circulating lane and one for the outer circulating lane) were considered. The Tanner bunching model was used for estimating φ parameter (Tanner, 1962). Right-lane capacity and left-lane capacity of entries 2-4 (see Figure 1), as well as right-lane capacity of entries 1-3 (see Figure 1) were estimated considering the circulating traffic flow in the outer lane ($Q_{c,e}$) at the subject entry approach from time to time considered:

$$
C_e = Q_{c,e} \cdot \left(1 - \frac{\Delta \cdot Q_{c,e}}{3600} \right) \cdot \frac{exp\left(\frac{-Q_{c,e}}{3600} \cdot \left(T_c - \Delta \right) \right)}{1 - exp\left(\frac{-Q_{c,e}}{3600} \cdot T_f \right)}
$$

(2)

Left-lane capacity estimations at minor entries 1-3 (see Figure 1) was estimated, instead, considering circulating traffic flows in the outer ($Q_{c,e}$) and in the inner lane ($Q_{c,i}$) on the circulatory carriageway:

$$
C_e = \left(Q_{c,e} + Q_{c,i} \right) \cdot \left(1 - \frac{\Delta \cdot Q_{c,e}}{3600} \right) \cdot \left(1 - \frac{\Delta \cdot Q_{c,i}}{3600} \right) \cdot \frac{exp\left(\frac{-Q_{c,e}}{3600} \cdot \left(T_{c,e} - \Delta \right) \frac{-Q_{c,i}}{3600} \cdot \left(T_{c,i} - \Delta \right) \right)}{1 - exp\left(\frac{-\left(Q_{c,e} + Q_{c,i} \right)}{3600} \cdot T_f \right)}
$$

(3)

Notations in Equations 2 and 3 have the same meaning as in Equation 1. It must be said that for each gap acceptance parameter a weighted mean was assumed starting from values surveyed by Fortuijn (2009b) and so specified:

-left entry lane at major entry (entry 2 or 4): $T_{c,e} = 3.60$ s, $T_f = 2.26$ s, $= 2.10$ s;

-right entry lane at major entry (entry 2 or 4): $T_{c,e}$ = 3.87 s, T_f = 2.13 s, = 2.10 s;

-left entry lane at minor entry (entry 1 or 3): $T_{c,i}$ = 3.19 s, $T_{c,e}$ = 3.03 s, T_f = 2.26 s, = 2.10 s;

-right entry lane at minor entry (entry 1 or 3): $T_{c,e}$ = 3.74 s, T_f = 2.13 s, = 2.10 s;

2.2 Capacity Models for Double-lane Roundabouts

Literature presents several operational models used for analysing performances at roundabouts. One of the first models was developed by Harders (1968); afterwards the same model was introduced into different edition of the Highway Capacity Manual (2000; 2010). Brilon et al. (1997) used the Tanner capacity equation (1962) for uncontrolled intersection adjusting it to needs of roundabout analysis. More recently, Brilon (2005) focused on the empirical regression of on-field experimental data and reached a simplified form of the capacity equation derived from the Siegloch's equation (1973); values for behavioral parameters were also proposed by Brilon (2005). Recent adaptations of the Siegloch's equation for capacity estimations of right and left entry lanes opposed by two conflicting lanes are reported in NCHRP 672 (2010).

A comprehensive summary of operational models can be found in the NCHRP 572 as drawn up by Rodegerdts et al. (2007). A recent estimation of gap acceptance parameters for roundabout capacity model applications is reported by Gazzarri et al. (2012).

2.2.1 Entry Capacity at Double-lane Roundabouts

This section focuses on assumptions adopted for evaluating entry capacities at double-lane roundabout in Figure 2. How to enter a roundabout is well known: entering vehicles face one or two circulating streams, depending on the entry lane by which they come from. It has been noted here that, although preferable, drivers do not have to preselect their entering lane in relation to their destination. So entry capacity estimations at the double-lane roundabouts under examination were obtained adding capacities of each entering lane. The shifted negative bunched exponential distribution (or Cowan's M3 headway distribution) was assumed to model circulating traffic flows; moreover, each entry lane capacity was calculated by using the Hagring model (1998) easily adapted to consider not only a single circulating stream (for estimating right-entry lane capacity by Equation 2), but also two circulating traffic streams (for estimating left-entry lane capacity by Equation 3).

In this manner, the circulating traffic flow was divided in the inner stream and the outer stream, the latter consisting in vehicles exiting from the intersection at the exit immediately after the considered entry approach. T_c and T_f were assumed equal to values reported in section 2.1.1 for right and left lanes at entries 1-3; this is due to manoeuvre schemas at a double-lane roundabout are considered analogous to those observed for the minor road of a turbo-roundabout. Δ was assumed equal to 2.10 s.

3. Comparing Basic Turbo-roundabouts and Double-lane Roundabouts to Evaluate Operational Benefits

In representing operating conditions at the intersections under examination, two traffic situations were analysed; the corresponding O-D matrices in percentage terms are reported in Table 1 (see case 1 and case 2). Assumptions concerned traffic demand: Q_{e2} was set equal to Q_{e4} and Q_{e1} was set equal to Q_{e3}; the cases with the overall entry flow coming from major entries ($Q_{e2}+Q_{e4}$) less than the overall entry flow coming from minor entries ($Q_{e3}+Q_{e1}$) were excluded.

With reference to the entry-lane selection performed by turning vehicles from entries, the following percentages were specified: i) at minor entries 1-3, right-turning vehicles were 90 percent from right-entry lane and 10 percent from left-entry lane; ii) at major entries 2-4, through vehicles were 50 percent both from right-entry lane and from left-entry lane.

Figure 3 shows the outcome of the comparison between the schemes in Figures 1 and 2 in terms of suitability domains obtained for the degree-of-saturation both with reference to the case 1 and the case 2 in Table 1 under undersaturation conditions.

In Figure 3a, corresponding to the case 1 in Table 1, it is possible to note the efficiency of double-lane roundabouts, performing better than turbo-roundabouts almost in the entire range of variation of entering traffic flows. In the Figure 3b it is possible to observe that turbo-roundabouts perform better than double-lane roundabouts when traffic flow coming from major roads maintain high levels; this condition occurs again when medium-to-high traffic flows enter the intersection from entry approaches.

Table 1. Origin/destination matrices of traffic flows in percentage terms

Case 1

O/D	1	2	3	4
1	0	0.33	0.33	0.33
2	0.33	0	0.33	0.33
3	0.33	0.33	0	0.33
4	0.33	0.33	0.33	0

Case 2

O/D	1	2	3	4
1	0	0.65	0.05	0.30
2	0.05	0	0.05	0.90
3	0.05	0.30	0	0.65
4	0.05	0.90	0.05	0

Description: o-d matrices of traffic flows in percentage terms representing flow scenarios, chosen to explore how different traffic patterns can influence operations. In the case 1 traffic flow percentages in o-d matrix were shared equally. In the case 2 percentages of through vehicles from and to major entries were considered significant compared to other turning vehicles; percentages of left and right turning vehicles from minor entries were significant compared to through vehicles from and to minor entries.
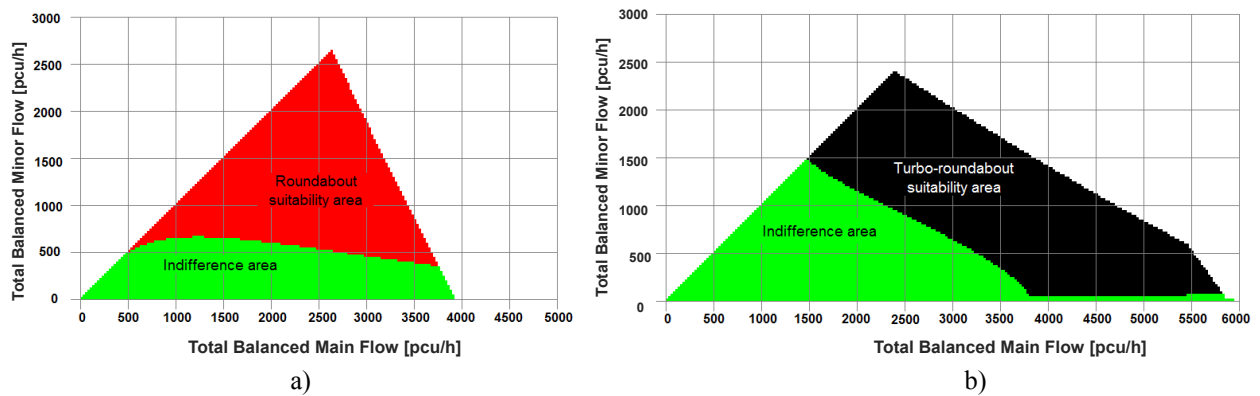

a)   b)

Figure 3. Suitability domains in undersaturated conditions in terms of degrees of saturation

Description of figure: it has be noted that x-axis represents the variable $(Q_{e2}+Q_{e4})$; the y-axis represents the variable $(Q_{e1}+Q_{e3})$; these variables are the basis for constructing suitability domains in undersaturated flow conditions having the following distinction between suitability areas:

- ■ degrees of saturation at roundabout less than 90 % than turbo-roundabout

- ■ degrees of saturation at turbo-roundabout less than 90 % than roundabout

- ■ indifference area

However, most appropriate details about the actual convenience of a pattern on the other can be derived from constructing analogous suitability domains in terms of delay experienced by users, considering the relation between the latter and the level of service quality. In order to perform the comparison between intersections under examination, the control delay at each intersection was computed as follows:

$$\bar{d} = \frac{\sum_{i=1}^{4} d_i \cdot Q_{e,i}}{\sum_{i=1}^{4} Q_{e,i}} \tag{4}$$

Symbols in the above equation require to be specified in relation to the two circular intersections here examined. In the case of the turbo-roundabout shown in Figure 1, the above parameter was calculated as the weighted mean value of the mean control delay d_i at each entering lane i, estimation of which starts from entry lane capacity $C_{e,i}$ and the degree of saturation. In the case of double-lane roundabout shown in Figure 2, d_i represents the control delay at each entry i, estimation of which starts from capacity of the entry approach in its entirety and the degree of saturation. With these specification d_i estimations were made using the analytical model given by HCM (2000) in chapter 17 to estimate the control delay at unsignalized intersections, which can be also used for roundabouts:

$$d_i = \frac{3600}{C_{e,i}} + 900T \left[\frac{Q_{e,i}}{C_{e,i}} - 1 + \sqrt{\left(\frac{Q_{e,i}}{C_{e,i}} - 1 \right)^2 + \frac{\left(\frac{3600}{C_{e,i}} \right)\left(\frac{Q_{e,i}}{C_{e,i}} \right)}{450T}} \right] + 5 \qquad (5)$$

where T is the length of analysis time period (h).

Figures 4 and 5 show the domains of convenience in undersaturated flow conditions obtained for control delay in the two situations corresponding to o-d matrices of traffic flows in percentage terms showed in Table 1. The distinction between suitability areas that can be identified in the two cases examined deserves an explanation. The light gray area in Figure 4 shows the situation of convenience for the double-lane roundabout; it represents, being equal entering traffic flows, the situation where delays experienced by users are less than 50% of those experienced by entering vehicles at turbo-roundabouts. The dark gray area in Figure 5 represents the situation of convenience for the turbo-roundabout, wherein delays are less than 50 percent of those experienced by entering vehicles at double-lane roundabouts. The situation in which delays in one of the two intersections are never less than 50% of those in which users can incur at the other intersection is showed in both figures by the area with a shade of gray intermediate between those used to identify the suitability area of the double-lane roundabout (see Figure 4) or the turbo-roundabout (see Figure 5).
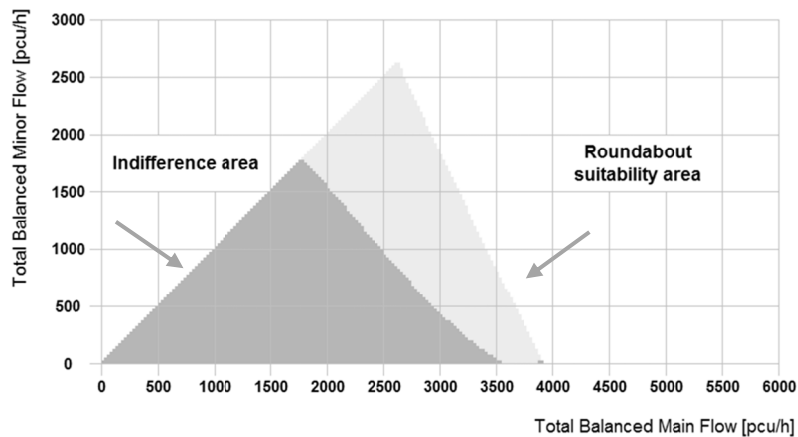

Figure 4. Example of suitability domain in undersaturated traffic conditions: case 1 in Table 1

Description: it has be noted that x-axis represents the variable $(Q_{e2}+Q_{e4})$; the y-axis represents the variable $(Q_{e1}+Q_{e3})$; these variables are the starting point for constructing suitability domains in undersaturated flow conditions having the following distinction:

	Roundabouts suitability area		Indifference area
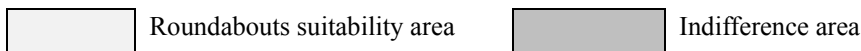
4. Results

It must be said that the convenience of the two intersections under examination was estimated in terms of delay experienced by entering vehicles, the latter being related to measures characterizing operational conditions or describing service quality at unsignalized intersections. Operating convenience of an intersection on the other, with reference to roundabouts here examined (see Figures 1 and 2) was obtained, indeed, preferring the delay experienced by users over the degree of saturation, measuring the former through the control delay.

In Figure 4 representing situations where traffic flow percentages were shared equally (see case 1 in Table 1), it is possible to observe the efficiency of double-lane roundabout compared to turbo-roundabout; it occurs when major entries are featured by high traffic flows entering the roundabout, and minor roads provide for low entering volumes. This condition still occurs when minor roads are featured by traffic flows growing from medium-to-high values.

In the case depicted in Figure 5 corresponding to case 2 in Table 1, turbo-roundabouts are featured by operating conditions advantageous compared to double-lane roundabouts. Low delays are experienced by users at turbo-roundabouts compared to double-lane roundabouts when high traffic volumes come from major roads and low-to-medium traffic flows come from minor roads. This traffic condition still happens when traffic flows

coming from minor road continue to grow and middle-to-high values of traffic flows enter the intersection from major roads. In both cases, it is possible to observe that when low-to medium traffic volumes enter the intersection the two intersections perform in an equivalent way.
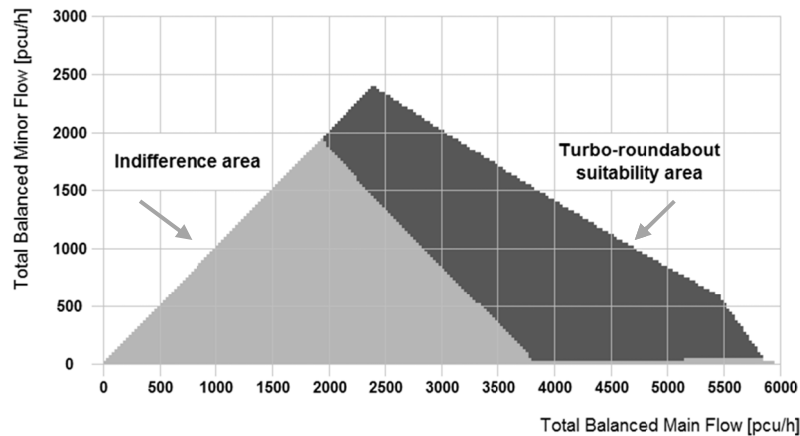

Figure 5. Example of suitability domain in undersaturated flow conditions: case 2 in Table 1

Description of figure: it has be noted that x-axis represents the variable $(Q_{e2}+Q_{e4})$; the y-axis represents the variable $(Q_{e1}+Q_{e3})$; these variables are the starting point for constructing suitability domains in undersaturated flow conditions having the following distinction:

turbo-roundabouts suitability area   indifference area

5. Conclusions

The primary purpose of this paper consisted in evaluating the convenience in terms of operational performances between basic turbo-roundabouts and double-lane roundabouts. Construction of suitability domains under undersaturation flow conditions (i.e., those conditions in which the demand volumes are less than approach capacity) was then reached for the two intersections under examination in each of the traffic situations explored.

This objective has been pursued making assumptions, found to be coherent by Authors, on the arrival process of circulating streams moving along the ring. Among entry capacity models based on the gap-acceptance theory here considered, applications of the Hagring model (1998) were proposed to evaluate entry capacity at intersections in Figures 1 and 2; for comparison purposes, the same model was specified for evaluating entry capacity at turbo-roundabouts and double-lane roundabouts. It must be said that the convenience of the two intersections under examination was estimated in terms of delay experienced by entering vehicles, considering the relation of this parameter to the level of service quality.

Results obtained for the two cases examined, despite they may be affected by the theoretical assumptions which the study takes as its starting point, showed that better performances at basic turbo-roundabouts than double-lane roundabouts are depending on the balance of entering traffic volumes at the approaches; they occur in particular when a significant share of traffic volumes is handled by major roads. Although there is a need of further cases to be examined in terms of traffic situations, the comparison developed in the paper, at last, can be helpful in choosing basic turbo-roundabout rather than double-lane roundabout and, in particular, to evaluate operating benefits obtainable from the conversion of an already existing double-lane roundabout to a turbo-roundabout with similar space requirements.

References

Al-Madani, H. M. N., & Saad, M. (2009). Analysis of roundabout capacity under high demand flows. In Urban Transport XV, *WIT Transactions on the Built Environment, 107*, 223-234.

Bovy, H., Dietrich, K., & Harmann, A. (1991). Guide Suisse des Giratoires. Lausanne, Switzerland, February 1991, p. 75 (cf. summary: Straße und Verkehr (Road and Traffic) No. 3, March 1991, pp. 137-139).

Brilon, W., Wu, N., & Bondzio, L. (1997). Unsignalized intersections in Germany. A state of the art 1997. Proceedings of the 3rd International Symposiu on Intersections without traffic signals (pp. 61-70), Portland, Oregon, USA, July, 1997.

Brilon, W., & Bäumer, H. (2004). Überprüfung von Kreisverkehren mit zweistreifig markierter oder einstreifg

markierter, aber zweistreifig befahrbarer Kreisfahrbahn. Forschung Straßenbau und Straßenverkehrstechnik, Heft 876, 2004.

Brilon, W. (2005). Roundabouts: A State of the Art in Germany. In Proceedings of the National Roundabout Conference 2005, Vail, Colorado, USA.

Brilon, W. (2008). *Turbo-Roundabout: an Experience from Germany*. [PowerPoint slides]. Retrieved from http://www.teachamerica.com/rab08/RAB08_Papers/RAB08S6BBrilon.pdf

Cowan, R. J. (1975). Useful headway model. *Transportation Research, 9*(6), 371-375. http://dx.doi.org/10.1016/0041-1647(75)90008-8

Cowan, R. J. (1987). An extension of Tanner's results on uncontrolled intersections. *Queueing Systems: Theory and Applications, 1*(3), 249-263.

Fortuijn, L. G. H., & Harte, V. F. (1997). Multilane roundabouts: an exploration of new forms (in Dutch), Verkeerskundige werkdagen 1997, CROW, Ede.

Fortuijn, L. G. H. (2009a). Turbo Roundabouts: Design Principles and Safety Performance. *Transportation Research Record, 2096*, 16-24. http://dx.doi.org/10.3141/2096-03

Fortuijn, L. G. H. (2009b). Turbo roundabouts. Estimation of capacity. *Transportation Research Record, 2130*, 83-92. http://dx.doi.org/10.3141/2130-11

Gazzarri, A., Martello, M. T., Pratelli, A., & Souleyrette, R. R. (2012). Estimation of gap acceptance parameters for HCM 2010 roundabout capacity model applications. Urban Transport XVIII, *WIT Transactions on The Built Environment, 128*, 309-320.

Giuffrè, O., Granà, A., Giuffrè, T., & Marino, R. (2008). How to derive the analytical capacity model for not-conventional urban roundabouts. Urban Transport XIV, *WIT Transactions on the Built Environment, 101*, 569-578. http://dx.doi.org/10.2495/UT080551

Giuffrè, O., Guerrieri, M., & Granà, A. (2009). Evaluating capacity and efficiency of turbo-roundabouts. Proceedings of the TRB 2009 Annual Meeting, Washington DC, USA.

Giuffrè, O, Granà, A., Giuffrè, T., & Marino, R. (2012). Researching a Capacity Model for Multilane Roundabouts with Negotiation of the Right-of-Way between Antagonist Traffic Flows. *Modern Applied Science, 6*(5), 2-12. http://dx.doi.org/10.5539/mas.v6n5p2

Hagring, O. (1998). A further generalization of Tanner's formula. *Transportation Research Part B: Methodological, 32*(6), 423-429. http://dx.doi.org/10.1016/S0191-2615(98)00010-1

Harders, J. (1968). The capacity of unsignalized urban intersections *[Die Leistungsfaehigkeit nicht signalgeregelter staedtischer Verkehrsknoten]* (in German). *Schriftenreihe Strassenbau und Strassenverkehrstechnik, 76*, 1968.

Luttinen, R. T. (1999). Properties of Cowan's M3 headway distribution. *Transportation Research Record, 1678*, 189-196. http://dx.doi.org/10.3141/1678-23

Mauro, R., & Branco, F. (2010). Comparative analysis of compact multilane roundabouts and turbo-roundabouts, *Journal of Transportation Engineering, 136*(4), 316-322. http://dx.doi.org/10.1061/(ASCE)TE.1943-5436.0000106

National Cooperative Highway Research Program (2010). Roundabouts: An Informational Guide (2nd ed.), Report 672 NCHRP, Washington DC, USA.

Pratelli, A., & Al-Madani, H. M. N. (2011). Testing for a large roundabouts capacity model: experimental comparisons between Italy and Bahrain. Urban Transport XVII. *WIT Transactions on the Built Environment, 116*, 3-15. http://dx.doi.org/10.2495/UT110011

Rodegerdts, L., Blogg, M., Wemple, E., Myers, E., Kyte, M., Dixon, M., … Carter, D. (2007). Roundabouts in the United States. Washington, D.C., USA: Transportation Research Board of the National Academies, NCHRP Report 572, 2007. Retrieved from http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_572.pdf

Siegloch, W. (1973). Capacity calculations for unsignalized intersection (in German). *Schriftenreihe Strassenbau und Strassenverkehrstechnik, 154*, 1973

Tanner, J. C. (1962). A theoretical analysis of delay At An Uncontrolled Intersections. *Biometrica, 49*(1-2), 163-170. http://dx.doi.org/10.1093/biomet/49.1-2.163

Tollazzi, T., Turnšek, S., & Renčelj, M. (2011). *Slovenian experiences with "Turbo-Roundabouts".* [PowerPoint slides]. Retrieved from http://teachamerica.com/RAB11/RAB1116Rencelj/player.html

Transportation Research Board. (2000, 2010). Highway Capacity Manual 2000, 2010, the 4th and the 5th edition of the HCM, edition 2000, 2010.

Troutbeck, R. J. (1986). Average Delay at an Unsignalized Intersection with Two Major Streams Each Having a Dichotomized Headway Distribution. *Transportation Science, 20*(4), 272-286. http://dx.doi.org/10.1287/trsc.20.4.272

Troutbeck, R. J. (1990). Traffic Interactions at Roundabouts. I*n Proceedings of the 15th Australian Road Research Board Conference, 15*(5), pp. 17-42, Darwin, Northern Territory, Australia.

Call for Manuscripts

Modern Applied Science (MAS) is an international, double-blind peer-reviewed, open-access journal, published by the Canadian Center of Science and Education. It publishes original research, applied, and educational articles in all areas of applied science. It provides an academic platform for professionals and researchers to contribute innovative work in the field. The scopes of the journal include, but are not limited to, the following fields: agricultural and biological engineering, applied mathematics and statistics, applied physics and engineering, chemistry and materials sciences, civil engineering and architecture, computer and information sciences, energy, environmental science and engineering, mechanics. The journal is published in both print and online versions. The online version is free access and download.

We are seeking submissions for forthcoming issues. All manuscripts should be written in English. Manuscripts from 3000–8000 words in length are preferred. All manuscripts should be prepared in MS-Word format, and submitted online, or sent to: mas@ccsenet.org

Paper Selection and Publishing Process

a) Upon receipt of a submission, the editor sends an e-mail of confirmation to the submission's author within one to three working days. If you fail to receive this confirmation, your submission e-mail may have been missed.

b) Peer review. We use a double-blind system for peer review; both reviewers' and authors' identities remain anonymous. The paper will be reviewed by at least two experts: one editorial staff member and at least one external reviewer. The review process may take two to four weeks.

c) Notification of the result of review by e-mail.

d) If the submission is accepted, the authors revise paper and pay the publication fee.

e) After publication, the corresponding author will receive two hard copies of the journal, free of charge. If you want to keep more copies, please contact the editor before making an order.

f) A PDF version of the journal is available for download on the journal's website, free of charge.

Requirements and Copyrights

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the authorities responsible where the work was carried out, and that, if accepted, the article will not be published elsewhere in the same form, in English or in any other language, without the written consent of the publisher. The editors reserve the right to edit or otherwise alter all contributions, but authors will receive proofs for approval before publication.

Copyrights for articles are retained by the authors, with first publication rights granted to the journal. The journal/publisher is not responsible for subsequent uses of the work. It is the author's responsibility to bring an infringement action if so desired by the author.

More Information

E-mail: mas@ccsenet.org

Website: www.ccsenet.org/mas

Paper Submission Guide: www.ccsenet.org/submission

Recruitment for Reviewers: www.ccsenet.org/reviewer

The journal is peer-reviewed
The journal is open-access to the full text
The journal is included in:

CABI
Chemical Abstracts database
DOAJ
EBSCOhost
Excellence in Research Australia (ERA)
Google Scholar

Open J-Gate
ProQuest
Scopus
Standard Periodical Directory
Ulrich's
Universe Digital Library